

Robust Phase Retrieval Algorithm for Time-Frequency Structured Measurements

Götz E. Pfander*

Palina Salanevich†

November 9, 2016

Abstract

We address the problem of signal reconstruction from intensity measurements with respect to a measurement frame. This non-convex inverse problem is known as *phase retrieval*. The case considered in this paper concerns phaseless measurements taken with respect to a Gabor frame. It arises naturally in many practical applications, such as diffraction imaging and speech recognition. We present a reconstruction algorithm that uses a nearly optimal number of phaseless time-frequency structured measurements and discuss its robustness in the case when the measurements are corrupted by noise. We show how geometric properties of the measurement frame are related to the robustness of the phaseless reconstruction. The presented algorithm is based on the idea of polarization as proposed by Alexeev, Bandeira, Fickus, and Mixon [1, 6].

Index terms: phase retrieval, Gabor frames, expander graphs, angular synchronization, spectral clustering, order statistics.

1 Introduction

The *phase retrieval* problem arises naturally in many applications within a variety of fields in science and engineering. Among these applications are optics [27], astronomical imaging [16], and microscopy [26].

As an example, let us consider the diffraction imaging problem [9]. To investigate the structure of a small particle, such as a DNA molecule, see [34], we illuminate the particle with X-rays and then measure the radiation scattered from it. When X-ray waves pass by an object and are measured in the far field, detectors are not able to capture the phase of the waves reaching them, but only their magnitudes. The measurements obtained in this way are of the form of pointwise squared absolute values of the Fourier transform of the object x , that is, the measurement map \mathcal{A} is given by $\mathcal{A}(x) = \{|\mathcal{F}(x)(n)|^2\}_{n \in \Omega}$ where Ω is the sampling grid. Since \mathcal{A} is not injective, some additional a priori information on the object x is needed for reconstruction.

*Department of Mathematics and Computer Science, Philipps-Universität Marburg, D-35032 Marburg, Germany. Email: pfander@mathematik.uni-marburg.de.

†Department of Mathematics and Logistics, Jacobs University Bremen, 28759 Bremen, Germany. Email: p.salanevich@jacobs-university.de.

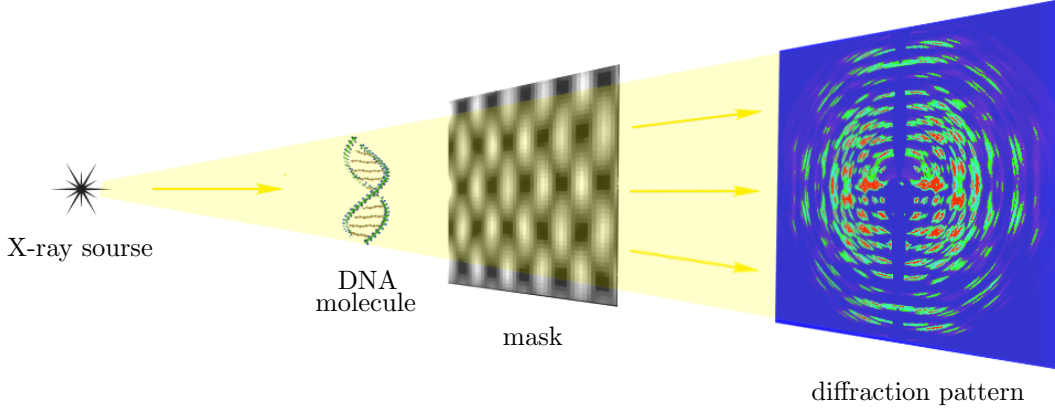


Figure 1: A typical setup for structured illuminations in diffraction imaging using a phase mask.

For instance, knowledge of the chemical interactions between parts of a DNA molecule were used for the construction of the DNA double helix model in the Nobel Prize winning work of Watson, Crick, and Wilkins [34].

One way to overcome this non-injectivity when no a priori information is available is *masking*. To modify the phase front, one can insert a known mask after the object, as shown on Figure 1. The measurement map in this case is given by $\mathcal{A}(x) = \{|\mathcal{F}(x \odot f_t)(n)|^2\}_{n \in \Omega, t \in I}$, where f_t , $t \in I$, are the masks used, and \odot denotes pointwise multiplication. By increasing the number of measurements in this way, we reduce the ambiguity in the reconstruction of x .

Since the problem of signal reconstruction from magnitudes of Fourier coefficients is particularly hard to handle, a more general frame theoretical setting is frequently considered. Namely, for $\Phi = \{\varphi_j\}_{j=1}^N \subset \mathbb{C}^M$ being a *frame*, that is, a spanning set for \mathbb{C}^M , we aim to recover $x \in \mathbb{C}^M$ from its phaseless squared frame coefficients $\mathcal{A}_\Phi(x) = \{|\langle x, \varphi_j \rangle|^2\}_{j=1}^N$.

Note that the masked Fourier coefficients of the signal x with masks $\{f_t\}_{t \in I} \subset \mathbb{C}^M$ can be viewed as the frame coefficients of x with respect to the frame Φ given by $\Phi = \{\varphi_{t,j}\}$, where $\varphi_{t,j}(m) = \frac{e^{2\pi i j m / M}}{\sqrt{M}} \bar{f}_t(m)$.

It is clear that, even in optimal setting, x can be reconstructed from intensity measurements only up to a global phase. Indeed, for every $\theta \in [0, 2\pi)$, the signals x and $e^{i\theta}x$ produce the same intensity measurements. Thus, the goal of phase retrieval is to reconstruct the equivalence class $[x] \in \mathbb{C}/\sim$ of x , where $x \sim y$ if and only if there exists $\theta \in [0, 2\pi)$, such that $x = e^{i\theta}y$. In the sequel, we are going to identify x with its equivalence class $[x] \in \mathbb{C}/\sim$.

Obviously, not every frame gives rise to an injective measurement map. But even in the case when \mathcal{A}_Φ is known to be injective, the problem of reconstructing $[x]$ from $\mathcal{A}_\Phi(x)$ is NP-hard in general [30]. So, the main goals in this area of applied mathematics is to find conditions on the number of measurements N and vectors φ_j for which there exist an efficient and robust numerical recovery algorithm.

Until recently, very little was known on how to achieve robust and efficient reconstruction given injectivity. Many practical methods used today have their origins in the alternating projection algorithms proposed in the 1970s by Gerchberg and Saxton [17]. Due to their lack of global convergence guarantees, the problem of developing fast phase retrieval algorithms which have provable recovery and robustness guarantees receives significant attention today. Some

of the most prominent suggested algorithms are PhaseLift [4, 10, 14, 11, 12], Wirtinger flow algorithm [13, 15], fast phase retrieval algorithm from local correlation measurements [20], and phase retrieval with polarization [1, 6]. The latter is described in more details in Section 2.1.

While recovery guarantees have been established for all these algorithms, most of them are designed to work exclusively when the frame vectors are independent random vectors. Since measurements of this type are not implementable in practice, the design of fast and stable recovery algorithms with a small number of application relevant, structured, measurements remains an important problem. We address this problem below.

1.1 Main result

We study the phase retrieval problem with measurement frame being a *Gabor frame*, that is, a collection of time and frequency shifts of a randomly chosen vector, called the *Gabor window* (see Section 2.3 for a precise definition). The main motivation for using Gabor frames is that in this case the frame coefficients are of the form of masked Fourier coefficients, where the masks are time shifts of the Gabor window. This makes measurements implementable in applications, but at the same time preserves the flexibility of the frame-theoretic approach. Apart from diffraction imaging mentioned above, phase retrieval with Gabor frames also arises, for example, in speech recognition problem [5].

Based on the idea of polarization [1], we propose a reconstruction algorithm for time-frequency structured measurements and investigate its robustness. In particular, we prove the following result, a more precise formulation of which we state in Section 4.2 as Theorem 4.7.

Theorem 1.1. *Fix $x \in \mathbb{C}^M$ and consider the time-frequency structured measurement frame constructed in (8). If the noise vector ν satisfies $\frac{\|\nu\|_2}{\|x\|_2} \leq \frac{c}{M}$ for some c sufficiently small, then there exists a numerical constant C , such that for the estimate \tilde{x} produced by Algorithm 5 the following holds with overwhelming probability*

$$\min_{\theta \in [0, 2\pi)} \|\tilde{x} - e^{i\theta} x\|_2^2 \leq C\sqrt{M}\|\nu\|_2.$$

We note that previous work on phase retrieval with Gabor frames concentrates on injectivity conditions for full Gabor frames and shows reconstruction from M^2 time-frequency structured measurements of an M -dimensional signal [8]. The reconstruction algorithm we propose in this paper requires only a close to optimal number $N = O(M \log M)$ of time-frequency structured measurements. To the best of our knowledge, Theorem 4.7 provides the best existing robustness guarantee for time-frequency structured measurements. Moreover, this result also guarantees exact (up to a global phase) recovery of x in the noiseless case, that is, when $\nu = 0$. Numerical experiments, presented in Section 5, verify the robustness of the proposed phase retrieval algorithm and illustrate dependencies of the error-to-noise ratio on various parameters.

The remaining part of this paper is organized as follows. In Section 2, we describe the idea of polarization and give some basic definitions and results from Gabor analysis and the theory of expander graphs that are used in the sequel. In Section 3, we describe the reconstruction algorithm for time-frequency structured measurements and discuss the robustness of this algorithm in Section 4. The analysis of robustness of the constructed algorithm leads to the investigation of geometric properties of Gabor frames, such as order statistics of the frame coefficients.

These properties are discussed in Section 4.1. Numerical results of the algorithm's robustness are presented in Section 5.

2 Notation and setup

Here and in the sequel, \odot denotes pointwise multiplication of two vectors of the same dimension. We view a vector $x \in \mathbb{C}^M$ as a function $x : \mathbb{Z}_M \rightarrow \mathbb{C}$, that is, all the operations on indices are done modulo M and $x(m - k) = x(M + m - k)$. We denote the complex unit sphere by $\mathbb{S}^{M-1} = \{x \in \mathbb{C}^M, \text{ s.t. } \|x\|_2 = 1\}$.

For a matrix $A \in \mathbb{C}^{k \times m}$, its adjoint matrix is denoted by $A^* \in \mathbb{C}^{m \times k}$, and the smallest singular value of A is denoted by $\sigma_{\min}(A)$. The Kronecker product of two matrices A and B is denoted by $A \otimes B$. Also, by a slight abuse of notation, we identify a frame $\Phi = \{\phi_j\}_{j=1}^N \subset \mathbb{C}^M$ with its synthesis matrix, having the frame vectors ϕ_j as columns. For any $V \subset \{1, \dots, N\}$, we set $\Phi_V = \{\phi_j\}_{j \in V}$.

We denote the Bernoulli distribution with success probability p by $B(1, p)$. Further, $\mathcal{N}(\mu, \sigma)$ denotes the Gaussian distribution with mean μ and variance σ^2 , and $\mathbb{CN}(\mu, \sigma)$ denotes the complex valued Gaussian distribution.

2.1 Phase retrieval with polarization

The polarization approach to phase retrieval can be described as follows [1]. Suppose $\Phi_V = \{\varphi_j\}_{j \in V} \subset \mathbb{C}^M$ is a measurement frame. We consider the phase retrieval problem

$$\begin{aligned} & \text{find} && x \\ & \text{subject to} && |\langle x, \varphi_j \rangle|^2 = b_j. \end{aligned} \quad (1)$$

For any $(i, j) \in V \times V$ with $|\langle x, \varphi_i \rangle| \neq 0$ and $|\langle x, \varphi_j \rangle| \neq 0$, we define the *relative phase* between frame coefficients as

$$\omega_{ij} = \left(\frac{\langle x, \varphi_i \rangle}{|\langle x, \varphi_i \rangle|} \right)^{-1} \frac{\langle x, \varphi_j \rangle}{|\langle x, \varphi_j \rangle|} = \frac{\overline{\langle x, \varphi_i \rangle} \langle x, \varphi_j \rangle}{|\langle x, \varphi_i \rangle| |\langle x, \varphi_j \rangle|}.$$

Note that $\omega_{ij}\omega_{jk} = \omega_{ik}$. Suppose that we are given $\{\omega_{ij}\}_{(i,j) \in E}$ for some set $E \subset V \times V$ in addition to the phaseless measurements with respect to Φ_V . Then we seek to solve the simpler problem

$$\begin{aligned} & \text{find} && x \\ & \text{subject to} && \frac{\overline{\langle x, \varphi_i \rangle} \langle x, \varphi_j \rangle}{|\langle x, \varphi_i \rangle| |\langle x, \varphi_j \rangle|} = \omega_{ij}, \\ & && |\langle x, \varphi_i \rangle|^2 = b_i. \end{aligned} \quad (2)$$

This problem can be solved using *phase propagation*. More precisely, we choose $|\langle x, \varphi_{i_0} \rangle| \neq 0$ and set

$c_{i_0} = |\langle x, \varphi_{i_0} \rangle|$. Then, for any $j \in V$ with $(i_0, j) \in E$, we define

$$c_j = \begin{cases} \omega_{i_0 j} |\langle x, \varphi_j \rangle| & \text{if } |\langle x, \varphi_j \rangle| \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

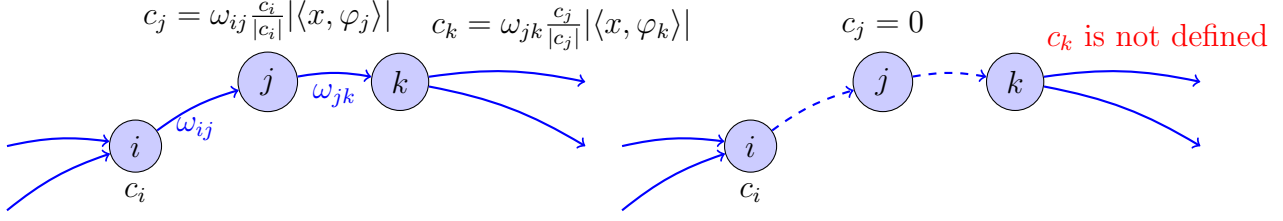


Figure 2: Phase propagation process (left). Phase propagation through vertex j fails due to zero measurement in this vertex (right).

In the next step, for each k with c_k not defined yet and $(i_0, j), (j, k) \in E$ for some j with $b_j \neq 0$, we set

$$c_k = \begin{cases} \omega_{jk} \frac{c_j}{|c_j|} |\langle x, \varphi_k \rangle| & \text{if } |\langle x, \varphi_k \rangle| \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

We repeat this step iteratively until values c_i are assigned to all indices $i \in V$ that can be reached from i_0 using edges from E . This process is illustrated in Figure 2 (left).

Assume that we were able to compute c_i for all indices $i \in V$. Then, using a dual frame $\tilde{\Phi}_V = \{\tilde{\varphi}_i\}_{i \in V}$ and treating c_i 's as frame coefficients, we reconstruct a representative of the “up-to-a-global-phase” equivalence class $[x]$ as

$$\sum_{j \in V} c_j \tilde{\varphi}_j = \sum_{j \in V} \omega_{i_0 j} |\langle x, \varphi_j \rangle| \tilde{\varphi}_j = \left(\frac{\langle x, \varphi_{i_0} \rangle}{|\langle x, \varphi_{i_0} \rangle|} \right)^{-1} \sum_{j \in V} \langle x, \varphi_j \rangle \tilde{\varphi}_j = \left(\frac{\langle x, \varphi_{i_0} \rangle}{|\langle x, \varphi_{i_0} \rangle|} \right)^{-1} x \in [x].$$

Let us consider the graph $G = (V, E)$, later called the *graph of measurements*, with the set of vertices indexed by V and the set of edges E . From the phase propagation procedure, it is apparent that if $\langle x, \varphi_j \rangle = 0$ for some $j \in V$, then the corresponding relative phases ω_{ji} are not defined for all $i \in V$ and the phase cannot be propagated through vertex j . This has the effect of deleting vertex j from G , see Figure 2 (right). If G remains connected after deleting all “zero” vertices, then, for every vertex i , there exists a path from i_0 to i , and c_i can be computed. This solves problem (2).

Thus, the initial phase retrieval problem (1) is reduced to the problem of finding relative phases between pairs of frame coefficients from a set E , such that the corresponding graph of measurements $G = (V, E)$ satisfies strong connectivity properties (as expander graphs do, see Section 2.2). To obtain the relative phase between frame coefficients, the following form of polarization identity will be used.

Lemma 2.1. [1] *Let $\omega = e^{2\pi i/3}$. If $\langle x, \varphi_i \rangle \neq 0$ and $\langle x, \varphi_j \rangle \neq 0$, then*

$$\omega_{ij} = \frac{1}{3|\langle x, \varphi_i \rangle||\langle x, \varphi_j \rangle|} \sum_{k=0}^2 \omega^k |\langle x, \varphi_i + \omega^k \varphi_j \rangle|^2.$$

In other words, to compute the relative phase ω_{ij} between the nonzero frame coefficients $\langle x, \varphi_i \rangle$ and $\langle x, \varphi_j \rangle$, we are using three additional phaseless measurements of x with respect to $\varphi_i + \varphi_j$, $\varphi_i + \omega \varphi_j$, and $\varphi_i + \omega^2 \varphi_j$. This means that reconstruction of x using phase propagation

involves only phaseless measurements, namely, phaseless measurements with respect to the union $\Phi_V \cup \Phi_E$, where Φ_V is a “vertex” frame and $\Phi_E = \{\varphi_i + \omega^k \varphi_j\}_{(i,j) \in E, k \in \{0,1,2\}}$. Note that $|\Phi_V \cup \Phi_E| = |V| + 3|E|$.

In [1], it is shown that in the noiseless case, one can perform phase retrieval with polarization using only $O(M)$ measurements. This algorithm is robust provided Φ_V consists of independent Gaussian vectors and the number of measurements is $O(M \log M)$. In [6], Bandeira, Chen, and Mixon adapt the polarization method to magnitude measurements of masked Fourier transforms of the signal. Using additive combinatorics tools, the authors show that the graph of measurements they are using for reconstruction is sufficiently connected provided that the total number of measurements is $O(M \log M)$. However, no stability results were obtained in [6] for the case of structured measurements.

In Section 3 we use the idea of polarization to build a recovery algorithm for time-frequency structured measurements and show reconstruction and stability guarantees for the constructed algorithm.

2.2 Expander graphs

Let us consider an undirected d -regular graph $G = (V, E)$, that is, a graph where each vertex is incident to exactly d edges. The number of vertices $|V|$ is denoted by n . For any $S, T \subset V$, we denote the set of edges between S and T by $E(S, T) = \{(u, v), \text{ s.t. } u \in S, v \in T, (u, v) \in E\}$.

Definition 2.2.

1. The *edge boundary* of a set $S \subset V$ is given by $\partial S = E(S, V \setminus S)$.
2. The *(edge) expansion ratio* of G is defined as

$$h(G) = \min_{|S| \leq \frac{n}{2}} \frac{|\partial S|}{|S|}.$$

3. A sequence of d -regular graphs $\{G_i\}_{i \in \mathbb{N}}$ of strictly increasing size $|V(G_i)|$ is a family of *expander graphs* if there exists $\varepsilon > 0$ such that $h(G_i) \geq \varepsilon$, for all $i \in \mathbb{N}$.

Some explicit constructions of families of expander graphs can be found in [19, 23, 35].

The *adjacency matrix* $A = A(G)$ of G with $|V(G)| = n$ is an $n \times n$ matrix whose (u, v) entry is equal to the number of edges in G connecting vertices u and v . Being real and symmetric, $A(G)$ has n real eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$. We refer to the set $\{\lambda_i\}_{i=1}^n$ of eigenvalues of $A(G)$ as the *spectrum* of the graph G . The spectrum encodes information about the connectivity of the graph. For example, G is connected if and only if $\lambda_1 > \lambda_2$. Also, the following theorem shows how the second eigenvalue is related to the expansion ratio of the graph [3, 19].

Theorem 2.3. (Cheeger inequality) *Let G be a d -regular graph with spectrum $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ and set $\lambda = \max\{|\lambda_2|, |\lambda_n|\}$. Then*

$$\frac{d - \lambda}{2} \leq h(G) \leq \sqrt{2d(d - \lambda)}.$$

The value $\text{spg}(G) = \frac{d-\lambda}{d}$ is also known as the *spectral gap* of G . Theorem 2.3 shows that a d -regular graph has an expansion ratio bounded away from zero if and only if its spectral gap is bounded away from zero. So, in order to construct graphs with big expansion ratio, one needs to construct graphs with big spectral gap.

To use graphs for phase retrieval with polarization, the following version of [18, Lemma 5.2.] is useful, see also [1].

Lemma 2.4. *Let G be a d -regular graph. For all $\varepsilon \leq \frac{\text{spg}(G)}{6}$, the graph obtained by removing any εn vertices from G has a connected component of size at least $\left(1 - \frac{2\varepsilon}{\text{spg}(G)}\right) n$.*

2.3 Gabor frames for \mathbb{C}^M

Let us begin by defining two families of unitary operators on \mathbb{C}^M , namely, cyclic shift operators and modulation operators.

Definition 2.5.

1. *Translation* (or *time shift*) by $k \in \mathbb{Z}_M$, is given by

$$T_k x = T_k(x(0), x(1), \dots, x(M-1)) = (x(m-k))_{m \in \mathbb{Z}_M}.$$

That is, T_k just permutes entries of x using k cyclic shifts.

2. *Modulation* (or *frequency shift*) by $\ell \in \mathbb{Z}_M$ is given by

$$M_\ell x = M_\ell(x(0), x(1), \dots, x(M-1)) = (e^{2\pi i \ell m / M} x(m))_{m \in \mathbb{Z}_M}.$$

That is, M_ℓ multiplies $x = x(\cdot)$ pointwise with the harmonic $e^{2\pi i \ell(\cdot)/M}$.

The superposition $\pi(k, \ell) = M_\ell T_k$ of translation by k and modulation by ℓ is a *time-frequency shift operator*.

Definition 2.6. For $g \in \mathbb{C}^M \setminus \{0\}$ and $\Lambda \subset \mathbb{Z}_M \times \mathbb{Z}_M$, the set of vectors

$$(g, \Lambda) = \{\pi(k, \ell)g\}_{(k, \ell) \in \Lambda}$$

is called the *Gabor system* generated by the *window* g and the set Λ . A Gabor system which spans \mathbb{C}^M is a frame and is referred to as a *Gabor frame*.

The *discrete Fourier transform* $\mathcal{F} : \mathbb{C}^M \rightarrow \mathbb{C}^M$ plays a fundamental role in Gabor analysis. It is given pointwise by

$$\mathcal{F}x(\ell) = \sum_{m \in \mathbb{Z}_M} x(m) e^{-2\pi i m \ell / M}, \quad \ell \in \mathbb{Z}_M.$$

The *short-time Fourier transform* (or *windowed Fourier transform*) $V_g : \mathbb{C}^M \rightarrow \mathbb{C}^{M \times M}$ with respect to the window $g \in \mathbb{C}^M \setminus \{0\}$ is given pointwise by

$$(V_g x)(k, \ell) = \langle x, \pi(k, \ell)g \rangle = \mathcal{F}(x \odot T_k \bar{g})(\ell), \quad k, \ell \in \mathbb{Z}_M. \quad (3)$$

Equality (3) indicates that the short-time Fourier transform on \mathbb{C}^M can be efficiently computed using the *fast Fourier transform* (FFT), an efficient algorithm to compute the discrete Fourier transform of a vector. Phase retrieval with time-frequency structured measurements benefits from this, as it reduces the run time of a recovery algorithm.

As we shall use polarization for phase retrieval, we would like to choose a window function g so that the frame (g, Λ) is a *full spark frame*, that is, so that for any subset $S \subset (g, \Lambda)$ of frame vectors with $|S| \geq M$, S spans \mathbb{C}^M [2]. Note that if the full Gabor system $(g, \mathbb{Z}_M \times \mathbb{Z}_M)$ is full spark, then so is (g, Λ) , for any $\Lambda \subset \mathbb{Z}_M \times \mathbb{Z}_M$.

The following result was shown for M prime in [22] and for M composite in [24].

Theorem 2.7. *Let M be a positive integer and let Λ be a subgroup of $\mathbb{Z}_M \times \mathbb{Z}_M$ with $|\Lambda| \geq M$. Then, for almost all windows ζ on the complex unit sphere $\mathbb{S}^{M-1} \subset \mathbb{C}^M$, (ζ, Λ) is a full spark frame.*

A more detailed description of Gabor frames in finite dimensions and their properties can be found in [29].

3 Phase retrieval from Gabor measurements

We now describe our design of a measurement frame and the reconstruction process in the noiseless case, which is also addressed in [31], and then discuss to robustness of the algorithm in the case when measurements are corrupted by noise. The analysis of the algorithm's robustness leads to the investigation of geometric properties of the measurement frame, such as frame bounds and flatness of the vector of frame coefficients. These properties are not only important for the problem at hand, but are of general interest in Gabor analysis.

3.1 Measurement process and frame construction

Consider the phase retrieval problem (1) with measurement frame

$$\Phi = \{\varphi_j\}_{j=1}^N = \Phi_V \cup \Phi_E \subset \mathbb{C}^M,$$

where Φ_V is a Gabor frame and Φ_E is a set of vectors corresponding to the additional edge measurements. As described in Section 2.1, phaseless measurements with respect to Φ_E are used to compute relative phases between frame coefficients.

Let us specify Φ_V and Φ_E now. For Φ_V we choose the Gabor frame

$$\begin{aligned} \Phi_V &= (g, \Lambda) \text{ with } \Lambda = F \times \mathbb{Z}_M, \ F \subset \mathbb{Z}_M, \ |F| = K, \\ \text{and } g &\in \mathbb{C}^M \text{ uniformly distributed on the unit sphere } \mathbb{S}^{M-1} \subset \mathbb{C}^M. \end{aligned} \tag{4}$$

The integer K is fixed and does not depend on the ambient dimension M . That is, we consider all frequency shifts and only a constant number of time shifts. As equation (3) indicates, our measurements are magnitudes of masked Fourier transform coefficients with the masks being $T_k \bar{g}$, $k \in F$.

We choose

$$\begin{aligned} \Phi_E &= \{\pi(\lambda_1)g + \omega^t \pi(\lambda_2)g, \text{ s.t. } (\lambda_1, \lambda_2) \in E, t \in \{0, 1, 2\}\} \text{ with } \omega = e^{2\pi i/3}, \\ E &= \{(\lambda_1, \lambda_2), \text{ s.t. } \lambda_1 = (k_1, \ell_1), \lambda_2 = (k_2, \ell_2), k_1, k_2 \in F, \ell_2 - \ell_1 \in C\} \subset \Lambda \times \Lambda, \\ \text{and } C &= D \cup (-D) \setminus \{0\} \subset \mathbb{Z}_M \text{ with } \mathbf{1}_D(m) \sim \text{i.i.d. } B\left(1, \frac{d \log M}{M}\right), \end{aligned} \quad (5)$$

where $d > 0$ is a parameter to be specified later. Then

$$\pi(\lambda_1)g + \omega^t \pi(\lambda_2)g = \pi(k_1, \ell_1)g + \omega^t \pi(k_2, \ell_2)g = p_{(\ell_2 - \ell_1)k_1 k_2}(t) \odot \pi(k_1, \ell_1)g,$$

where the vector $p_{ck_1 k_2}(t) \in \mathbb{C}^M$ is defined pointwise by

$$p_{c, k_1, k_2}(t)(m) = 1 + e^{2\pi i(\frac{cm}{M} + \frac{t}{3})} \frac{g(m - k_2)}{g(m - k_1)}, \quad m \in \mathbb{Z}_M,$$

with parameters $c \in C$, $k_1, k_2 \in F$, and $t \in \{0, 1, 2\}$. Therefore, for each fixed parameter triple (c, k_1, k_2) , the respective additional measurements are of the form of magnitudes of masked Fourier transform coefficients as well, namely,

$$\{|\langle x, p_{ck_1 k_2}(t) \odot \pi(k_1, \ell)g \rangle|\}_{\ell \in \mathbb{Z}_M, t \in \{0, 1, 2\}} = \{|\mathcal{F}(x \odot \bar{p}_{ck_1 k_2}(t) \odot T_{k_1} \bar{g})(\ell)|\}_{\ell \in \mathbb{Z}_M, t \in \{0, 1, 2\}}. \quad (6)$$

Let us note that the frame $\Phi = \Phi_V \cup \Phi_E$ constructed in this way consists of $|\Lambda| + 3|E| = |F|M + 3|F|^2|C|M$ vectors. Since $C = D \cup (-D) \setminus \{0\}$ with $\mathbf{1}_D(m) \sim \text{i.i.d. } B(1, \frac{d \log M}{M})$, we have $|C| = O(\log M)$ with high probability and thus $|\Phi| = O(M \log M)$.

Using the polarization identity in Lemma 2.1 with $\omega = e^{2\pi i/3}$, we compute the relative phases

$$\omega_{\lambda_1 \lambda_2} = \left(\frac{\langle x, \pi(\lambda_1)g \rangle}{|\langle x, \pi(\lambda_1)g \rangle|} \right)^{-1} \frac{\langle x, \pi(\lambda_2)g \rangle}{|\langle x, \pi(\lambda_2)g \rangle|} = \frac{1}{3|\langle x, \pi(\lambda_1)g \rangle||\langle x, \pi(\lambda_2)g \rangle|} \sum_{t=0}^2 \omega^t |\langle x, \pi(\lambda_1)g + \omega^t \pi(\lambda_2)g \rangle|^2, \quad (7)$$

for $(\lambda_1, \lambda_2) \in E$, where E is defined by (5). Recall, that $\omega_{\lambda_1 \lambda_2}$ is well defined if and only if $|\langle x, \pi(\lambda_1)g \rangle| \neq 0$ and $|\langle x, \pi(\lambda_2)g \rangle| \neq 0$.

Remark 3.1. As equations (3) and (6) indicate, all required measurements are magnitudes of masked Fourier transform coefficients. These are relevant for many applications. Moreover, measurements and reconstruction in this case can be implemented using FFT, which allows a noticeable speed up of measurement and reconstruction processes. For comparison, the computational complexity of the measurement process with random Gaussian frame of cardinality $O(M \log M)$ (as considered, for example, in [14] and [1]) is $O(M^2 \log M)$, and the complexity of measurement with the frame Φ constructed above is $O(M \log^2 M)$. Furthermore, in the case of random Gaussian frames, we have to use $O(M^2 \log M)$ memory bits to store the measurement matrix, while for our frame Φ it is enough to store the window g and the set C , and the overall amount of memory used is only $M + O(\log M) = O(M)$. These are some of the advantages of time-frequency structured frames in comparison to randomly generated frames.

3.2 Reconstruction in the noiseless case

We now describe our polarization based reconstruction algorithm for time-frequency structures measurements.

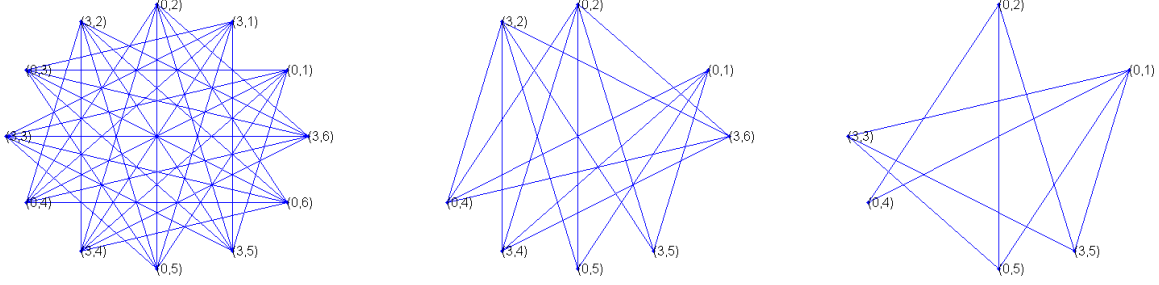


Figure 3: An example of the graph of measurements G with $M = 6$, $F = \{0, 3\}$ and $C = \{2, 3, 4\}$ (left). This graph remains connected after deleting one third of its vertices (middle). After we delete one half of its vertices it has a connected component of size at least 4 (right).

ALGORITHM 1 (RECONSTRUCTION IN THE NOISELESS CASE)

Input: Phaseless measurements b w.r.t. $\Phi_\Lambda \cup \Phi_E$, defined by (4) and (5);
 $F < \mathbb{Z}_M$, $C \subset \mathbb{Z}_M$, and window $g \in \mathbb{C}^M \setminus \{0\}$.

Reconstruction: Construct the graph $G = (\Lambda, E)$ with $\Lambda = F \times \mathbb{Z}_M$ and E as in (5);
assign to each $\lambda \in \Lambda$ the weight b_λ ;
assign to each edge $(\lambda_1, \lambda_2) \in E$ the weight $\omega_{\lambda_1 \lambda_2}$ computed using (7);
delete from G all vertices λ with $b_\lambda = 0$ to obtain $G' = (\Lambda', E') \subset G$;
choose a connected component $G'' = (\Lambda'', E'') \subset G'$ of the biggest size;
run the phase propagation process (Section 2.1) to obtain c_λ , $\lambda \in \Lambda''$;
reconstruct $\tilde{x} = (\Phi_{\Lambda''} \Phi_{\Lambda''}^*)^{-1} \Phi_{\Lambda''} c$ from $c = \{c_\lambda\}_{\lambda \in \Lambda''}$.

Output: $\tilde{x} \in [x]$, initial signal up to a global phase.

Since $0 \notin C$, the graph G constructed above has no loops, and since $C = -C$, it is not directed. Also, each vertex $\lambda = (k, \ell)$ of G is adjacent to any vertex $\lambda' = (k', \ell + c)$ with $c \in C$ and $k' \in F$. Thus, each vertex in G has degree $|F||C|$ and G is regular.

To be able to reconstruct a signal x using Algorithm 1, we need to ensure that $|\Lambda''|$ is sufficiently large, so that $\Phi_{\Lambda''} = (g, \Lambda'')$ is a frame. Then x can be recovered from its frame coefficients with respect to $\Phi_{\Lambda''}$. In terms of the graph of measurements G , this means that after we delete all vertices λ with zero weight, the resulting graph has a connected component of sufficiently big size. As Lemma 2.4 shows, this is satisfied provided that G has a sufficiently big spectral gap. The spectral gap of $G = (\Lambda, E)$ can be estimated in terms of the set C . More precisely, it depends on the Fourier bias of C , which is defined in [33].

Definition 3.2. Take $C \subset \mathbb{Z}_M$ and let $\mathbf{1}_C$ be the characteristic function of C . Then the *Fourier bias* of C is given by

$$\|C\|_u = \max_{m \neq 0} |(\mathcal{F}\mathbf{1}_C)(m)|.$$

In additive combinatorics the Fourier bias is used to measure pseudorandomness of a set. Roughly speaking, it helps to distinguish between sets which are highly uniform and behave like random sets, and those which are highly non-uniform and behave like arithmetic progressions. The following lemma relates the spectral gap of G and the Fourier bias $\|C\|_u$ of C [6, 33].

Lemma 3.3. *Consider the graph G constructed as above. Then the spectral gap of G satisfies*

$$\text{spg}(G) = 1 - \frac{\|C\|_u}{|C|}.$$

Proof. Consider the adjacency matrix $A = \{a_{(k-1)|F|+\ell, (k'-1)|F|+\ell'}\}_{(k,\ell), (k',\ell') \in \Lambda}$ of the graph G . For any fixed pair $k_1, k_2 \in F$ we have

$$a_{(k_1-1)|F|+\ell_1, (k_2-1)|F|+\ell_2} = \begin{cases} 0, & \text{if } \ell_1 - \ell_2 \in C; \\ 1, & \text{if } \ell_1 - \ell_2 \notin C. \end{cases}$$

It follows that $A = J \otimes \text{circ}(\mathbf{1}_C)$, where J is the $|F| \times |F|$ matrix with all entries being 1, and $\text{circ}(\mathbf{1}_C)$ is the circulant matrix generated by the vector $\mathbf{1}_C$, that is

$$\text{circ}(\mathbf{1}_C) = \begin{pmatrix} \mathbf{1}_C(0) & \mathbf{1}_C(M-1) & \cdots & \mathbf{1}_C(1) \\ \mathbf{1}_C(1) & \mathbf{1}_C(0) & \cdots & \mathbf{1}_C(2) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{1}_C(M-1) & \mathbf{1}_C(M-2) & \cdots & \mathbf{1}_C(0) \end{pmatrix}.$$

The eigenvalues of the Kronecker product matrix A are given by the pairwise products of eigenvalues of the factors

$$\lambda_{j,m}(A) = \lambda_j(J) \lambda_m(\text{circ}(\mathbf{1}_C)), \quad j \in \{1, \dots, |F|\}, \quad m \in \mathbb{Z}_M.$$

The only nontrivial eigenvalue of J is $\lambda_1(J) = |F|$ and $\lambda_m(\text{circ}(\mathbf{1}_C)) = \mathcal{F}(\mathbf{1}_C)(m)$, for each $m \in \mathbb{Z}_M$. Thus, for any $j > 1$ we have $\lambda_{j,m} = 0$ and the nontrivial eigenvalues of A are given by $\lambda_{1,m}(A) = |F| \sum_{k \in C} e^{-2\pi i k m / M}$. We have

$$|\lambda_{1,m}(A)| = |F| \left| \sum_{j \in C} e^{-2\pi i j m / M} \right| \leq |F| \|C\|_u,$$

with equality if and only if $m = 0$. Thus the spectral gap of G is

$$\text{spg}(G) = \frac{1}{|F| \|C\|_u} \left(|F| \|C\|_u - \max_{m \neq 0} |F| |(\mathcal{F}(\mathbf{1}_C))(m)| \right) = 1 - \frac{\|C\|_u}{|C|}.$$

□

The following result is shown in [6]. The proof is based on a version of the Chernoff bound and Lemma 3.3.

Lemma 3.4. *Pick $d > 36$ and suppose the entries of the characteristic vector $\mathbf{1}_D$ of a set D are independent with distribution $B\left(1, \frac{d \log M}{M}\right)$. Take $C = D \cup (-D) \setminus \{0\}$ and construct the graph G as above. Then with overwhelming probability*

$$\text{spg}(G) \geq 1 - \frac{6}{\sqrt{d}}.$$

Now we are ready to establish recovery guarantees for Algorithm 1.

Theorem 3.5. *Let frames Φ_V and Φ_E be constructed as above, with $|F| = 12$ and $d = 144$. Then every signal $x \in \mathbb{C}^M$ can be reconstructed from $M + 3|F|^2 M|C| = O(M \log M)$ phaseless measurements with respect to the frame $\Phi_V \cup \Phi_E$ using Algorithm 1.*

Proof. We begin the reconstruction algorithm with assigning to each vertex $\lambda \in \Lambda$ of the constructed graph G the weight $b_\lambda = |\langle x, \pi(\lambda)g \rangle|$ and assigning to each edge $(\lambda_1, \lambda_2) \in E$ of G the relative phase $\omega_{\lambda_1 \lambda_2}$ which is computed from the additional edge measurements. Theorem 2.7 implies that the frame $\Phi_V = \{\pi(\lambda)g\}_{\lambda \in \Lambda}$ is full spark with probability 1. Thus, for any vector $x \in \mathbb{C}^M$, the number of zero measurements among $\{b_\lambda = |\langle x, \pi(\lambda)g \rangle|\}_{\lambda \in \Lambda}$ is at most $M - 1$. In other words, Algorithm 1 deletes at most $M - 1$ vertices from G to obtain G' .

Next, Φ_V being full spark implies that any its subset $\Phi_{V'} \subset \Phi_V$ of size $|\Phi_{V'}| \geq M$ form a frame. Thus, to recover x , it is enough to know any M of the frame coefficients with respect to Φ_V . To show that G' has a connected component G'' of size at least M , first note that Lemma 3.4 ensures that $\text{spg}(G) \geq 1 - \frac{6}{\sqrt{d}} = \frac{1}{2}$. Then, applying Lemma 2.4 with $n = |\Lambda| = |F|M$ and $\varepsilon = \frac{1}{|F|} = \frac{1}{12} \leq \frac{\text{spg } G}{6}$, we obtain that after deleting any $\varepsilon n = M$ vertices from G , the largest connected component G'' will have at least $\left(1 - \frac{2\varepsilon}{\text{spg } G}\right)n \geq \frac{2}{3}|F|M = 8M > M$ vertices.

By running the phase propagation algorithm on the connected graph $G'' = (\Lambda'', E'')$, we recover c_λ , $\lambda \in \Lambda''$, which are, up to a global phase $e^{i\theta}$, $\theta \in [0, 2\pi)$, equal to the corresponding frame coefficients of x with respect to $\Phi_{\Lambda''}$, that is

$$c_\lambda = e^{i\theta} \langle x, \pi(\lambda)g \rangle, \quad \lambda \in \Lambda''.$$

Now, using the canonical dual frame, we obtain

$$\tilde{x} = (\Phi_{\Lambda''} \Phi_{\Lambda''}^*)^{-1} \Phi_{\Lambda''} c = e^{i\theta} (\Phi_{\Lambda''} \Phi_{\Lambda''}^*)^{-1} \Phi_{\Lambda''} \Phi_{\Lambda''}^* x = e^{i\theta} x.$$

□

4 Robustness of reconstruction in the presence of noise

In many applications, measurements are corrupted by noise. In this section we address the behaviour of the presented algorithm in the case when the available measurements are of the form

$$\begin{aligned} b_\lambda &= |\langle x, \pi(\lambda)g \rangle|^2 + \nu_\lambda, \quad \lambda \in \Lambda; \\ b_{\lambda_1 \lambda_2 t} &= |\langle x, \pi(\lambda_1)g + \omega^t \pi(\lambda_2)g \rangle|^2 + \nu_{\lambda_1 \lambda_2 t}, \quad (\lambda_1, \lambda_2) \in E, \quad t \in \{0, 1, 2\}, \end{aligned} \tag{8}$$

where ν_λ , $\nu_{\lambda_1 \lambda_2 t}$ are noise terms. We aim to construct a modification of Algorithm 1 which, in presence of noise, recovers a close estimate \tilde{x} of the original signal x .

4.1 Order statistics of frame coefficients

To compute a relative phase between two frame coefficients, we rely on formula (7). The calculations include division by $|\langle x, \pi(\lambda_1)g \rangle|$ and $|\langle x, \pi(\lambda_2)g \rangle|$ and are therefore very sensitive to perturbations when $|\langle x, \pi(\lambda_1)g \rangle|$ or $|\langle x, \pi(\lambda_2)g \rangle|$ is small. While “zero” vertices provide no relative phase information, vertices with small vertex measurements lead to unreliable relative phase estimations and should therefore be deleted from the graph. As we require the graph of measurements to have a connected component of size at least M after deleting vertices with small weights, the number of such vertices has to be estimated. To do so, we show that the vector of frame coefficients of a fixed $x \in \mathbb{S}^{M-1}$ with respect to a Gabor frame with random window is “flat” with high probability. That is, most of the frame coefficients are in the range of $\frac{c}{\sqrt{M}}$ to $\frac{K}{\sqrt{M}}$, for some suitably chosen constants $K > c > 0$.

Theorem 4.1. *Fix $x \in \mathbb{S}^{M-1} \subset \mathbb{C}^M$ and consider a Gabor frame (g, Λ) with $\Lambda \subset \mathbb{Z}_M \times \mathbb{Z}_M$ and a random window g uniformly distributed on the unit sphere \mathbb{S}^{M-1} . Then the following holds.*

(a) *For any $c > 0$ and $k > 0$, with probability at least $1 - \frac{1}{k^2}$, we have*

$$\left| \left\{ \lambda \in \Lambda, \text{ s.t. } |\langle x, \pi(\lambda)g \rangle| < \frac{c}{\sqrt{M}} \right\} \right| < |\Lambda|(c^2 + kc).$$

(b) *For any $K > 0$ and $k > 0$, with probability at least $1 - \frac{1}{k^2}$, we have*

$$\left| \left\{ \lambda \in \Lambda, \text{ s.t. } |\langle x, \pi(\lambda)g \rangle| > \frac{K}{\sqrt{M}} \right\} \right| < |\Lambda| \left(\frac{8}{\pi} e^{-K^2} + k \frac{2\sqrt{2}}{\sqrt{\pi}} e^{-\frac{K^2}{2}} \right).$$

The proof of this result is presented in the appendix.

Remark 4.2. A similar result can be shown for a Gabor frame with a window whose entries are independent Gaussian random variables. The proof in this case involves the same steps as the proof of Theorem 4.1.

Theorem 4.1 is a non-uniform result in the sense that the proven bounds hold with high probability for each *individual* x . Note that this does not imply that the same bounds will hold *simultaneously for all* $x \in \mathbb{S}^{M-1}$ with high probability. We give a uniform bound for the number of large frame coefficients in the following result.

Theorem 4.3. *Consider a Gabor frame $(g, \mathbb{Z}_M \times \mathbb{Z}_M)$ with window g whose entries are independent Gaussian random variables with zero mean and variance $\frac{1}{M}$. Then, for some suitably chosen numerical constants $c, c_1 > 0$,*

$$\mathbb{P} \left\{ \left| \left\{ \lambda \in \mathbb{Z}_M \times \mathbb{Z}_M, \text{ s.t. } |\langle x, \pi(\lambda)g \rangle| > \sqrt{\frac{3}{2c} \frac{\log^2 M}{M}} \right\} \right| < \frac{cM}{\log^4 M}, \text{ for all } x \in \mathbb{S}^{M-1} \right\} \geq 1 - e^{-c_1 \log^3 M}.$$

Proof. Let g be a random vector with $g(m) \sim \text{i.i.d. } \mathcal{CN}\left(0, \frac{1}{M}\right)$, $m \in \mathbb{Z}_M$, and let $\Phi = \Phi_{\mathbb{Z}_M \times \mathbb{Z}_M}$ be an $M \times M^2$ matrix whose columns are $\pi(\lambda)g$, $\lambda \in \mathbb{Z}_M \times \mathbb{Z}_M$. Fix $s \in \{1, \dots, M^2\}$ and for any $x \in \mathbb{S}^{M-1}$ denote by S_x the set of $\lambda \in \mathbb{Z}_M \times \mathbb{Z}_M$ corresponding to the s biggest in modulus

frame coefficients of x with respect to the Gabor frame $(g, \mathbb{Z}_M \times \mathbb{Z}_M)$. Then, for the phase vector $v_x \in \mathbb{C}^{M^2}$ defined by

$$v_x(\lambda) = \begin{cases} \frac{\langle x, \pi(\lambda)g \rangle}{|\langle x, \pi(\lambda)g \rangle|}, & \lambda \in S_x \\ 0, & \text{otherwise,} \end{cases}$$

we have

$$x^* \Phi v_x = \sum_{\lambda \in S_x} |\langle x, \pi(\lambda)g \rangle|.$$

Applying the Cauchy–Schwarz inequality to $x^* \Phi v_x = \langle \Phi v_x, x \rangle$, we obtain

$$\sum_{\lambda \in S_x} |\langle x, \pi(\lambda)g \rangle| \leq \|x\|_2 \|\Phi v_x\|_2 = \|\Phi v_x\|_2.$$

Note that v_x is an s -sparse vector with $\|v_x\|_2 = \sqrt{s}$. Then, if $s = \frac{cM}{\log^4 M}$ for a suitably chosen numerical constant $c > 0$, we have

$$\frac{1}{2} \|v\|_2^2 \leq \|\Phi v\|_2^2 \leq \frac{3}{2} \|v\|_2^2,$$

for any s -sparse vector $v \in \mathbb{C}^{M^2}$ with probability at least $1 - e^{-c_1 \log^3 M}$, where $c_1 > 0$ depends only on c , see [21, Theorem 5.1]. Thus

$$\sum_{\lambda \in S_x} |\langle x, \pi(\lambda)g \rangle| \leq \|\Phi v_x\|_2 \leq \sqrt{\frac{3s}{2}}$$

with probability at least $1 - e^{-c_1 \log^3 M}$. It follows that with the same probability,

$$\min_{\lambda \in S_x} |\langle x, \pi(\lambda)g \rangle| \leq \sqrt{\frac{3}{2s}} = \sqrt{\frac{3}{2c}} \frac{\log^2 M}{\sqrt{M}}.$$

In other words, with probability at least $1 - e^{-c_1 \log^3 M}$, for any $x \in \mathbb{S}^{M-1}$ all except at most $\frac{cM}{\log^4 M} - 1$ frame coefficients are in modulus bigger than $\sqrt{\frac{3}{2c}} \frac{\log^2 M}{\sqrt{M}}$. \square

Since Theorem 4.3 holds for the full Gabor frame $(g, \mathbb{Z}_M \times \mathbb{Z}_M)$, it also holds for any its subframe (g, Λ) , $\Lambda \subset \mathbb{Z}_M$. Also, note that while Theorem 4.3 gives a better bound on the number of large frame coefficients than Theorem 4.1 (b), it gives a slightly weaker bound on the modulus of the remaining coefficients, namely, $\frac{C \log^2 M}{\sqrt{M}}$ instead of $\frac{C}{\sqrt{M}}$.

Remark 4.4. Note that Theorem 5.1 in [21] is formulated for a window g with independent mean-zero, variance one, L -subgaussian entries. Thus, Theorem 4.3 is also true in this more general case.

4.2 Reconstruction from noisy measurements

To obtain a modification to Algorithm 1 that leads to robust reconstruction, we may assume, without loss of generality, that the signal x lies on the complex unit sphere $\mathbb{S}^{M-1} \subset \mathbb{C}^M$. As mentioned before, instead of deleting vertices with zero weight, a portion of vertices with small weights should be deleted in the first step of the reconstruction algorithm. As shown later, very large measurements can also prevent stable reconstruction, so we delete respective vertices as well. To delete vertices from the graph of measurements, we use the following algorithm.

ALGORITHM 2 (DELETING “SMALL” AND “LARGE” VERTICES)

Input: Graph $G = (\Lambda, E)$ with weighted vertices V , parameters α, β .
For $i = 0$ **to** $(1 - \alpha)|\Lambda|$ Find $\lambda \in \Lambda$ with the smallest value of b_λ ;
delete the vertex λ from G .
For $i = 0$ **to** $(1 - \beta)|\Lambda|$ Find $\lambda \in \Lambda$ with the largest value of b_λ ;
delete the vertex λ from G .
Output: Graph G' with more “flat” vertex weights.

Noise might accumulate while passing from one vertex to another in the phase propagation process. Thus, we seek an efficient method to reconstruct the phases of the vertex frame coefficients using measured relative phases. For this purpose, we shall use the *angular synchronisation algorithm* (Algorithm 4 below) whose main idea we describe now [32, 1].

Let $G'' = (\Lambda'', E'')$ be a subgraph of G' and let A be the weighted adjacency matrix of the graph G'' given by

$$A(\lambda_1, \lambda_2) = \begin{cases} \frac{\overline{\langle x, \pi(\lambda_1)g \rangle} \langle x, \pi(\lambda_2)g \rangle + \varepsilon_{\lambda_1 \lambda_2}}{|\overline{\langle x, \pi(\lambda_1)g \rangle} \langle x, \pi(\lambda_2)g \rangle + \varepsilon_{\lambda_1 \lambda_2}|}, & (\lambda_1, \lambda_2) \in E'' \\ 0, & (\lambda_1, \lambda_2) \notin E'', \end{cases} \quad (9)$$

where $\lambda_1, \lambda_2 \in \Lambda''$ and $\varepsilon_{\lambda_1 \lambda_2} = \frac{1}{3} \sum_{t=0}^2 \omega^t \nu_{\lambda_1 \lambda_2 t}$. Let $v_{\lambda_i} = \frac{\langle x, \pi(\lambda_i)g \rangle}{|\langle x, \pi(\lambda_i)g \rangle|}$, $\lambda_i \in \Lambda''$, be the true phases of the frame coefficients. Then $A(\lambda_i, \lambda_j)$ can be considered as an approximation of the relative phase $\omega_{\lambda_i \lambda_j} = v_{\lambda_i}^{-1} v_{\lambda_j}$. The vector $\tilde{v} = \{\tilde{v}_\lambda\}_{\lambda \in \Lambda''}$ given by

$$\tilde{v} = \arg \min_{|u_\lambda|=1, \lambda \in \Lambda''} \sum_{\{\lambda_i, \lambda_j\} \in E''} |u_{\lambda_j} - A(\lambda_i, \lambda_j) u_{\lambda_i}|^2 = \arg \min_{|u_\lambda|=1, \lambda \in \Lambda''} u^* (D - \bar{A}) u, \quad (10)$$

where D is the diagonal matrix of vertex degrees and \bar{A} is a componentwise conjugate of A . The vector \tilde{v} is an approximation of v . Note that we assume $|u_\lambda| = 1$, $\lambda \in \Lambda''$, and thus the quantity $u^* D u = \sum_{\lambda \in \Lambda''} d_\lambda |u_\lambda|^2 = \sum_{\lambda \in \Lambda''} d_\lambda$ does not vary with u . Minimizing the above quantity is equivalent to minimizing

$$\frac{u^* (D - \bar{A}) u}{u^* D u} = \frac{(D^{1/2} u)^* (I - D^{-1/2} \bar{A} D^{-1/2}) (D^{1/2} u)}{\|D^{1/2} u\|_2^2},$$

which is larger than or equal to the smallest eigenvalue α_1 of the *connected Laplacian* matrix $L_1 = I - D^{-1/2}\bar{A}D^{-1/2}$. Equality is achieved if $D^{1/2}u$ is a corresponding eigenvector.

Note that in the case when G'' is connected and no noise is present, (10) has a unique (up to a global phase) minimizer. The following result shows stability of the Algorithm 4. It is proven in [1, 7].

Theorem 4.5. *Consider a graph $G = (\Lambda'', E'')$ with spectral gap $\tau > 0$, and define $\|\theta\|_{\mathbb{T}} = \min_{k \in \mathbb{Z}} |\theta - 2\pi k|$ for all angles $\theta \in \mathbb{R}/2\pi\mathbb{Z}$. Then, given the weighted adjacency matrix A as in (9), Algorithm 4 outputs $\tilde{v} \in \mathbb{C}^{|\Lambda''|}$ with unit-modulus entries, such that for some phase $\theta \in \mathbb{R}/2\pi\mathbb{Z}$,*

$$\sum_{\lambda \in \Lambda''} \|\arg(\tilde{v}_\lambda) - \arg(\langle x, \pi(\lambda)g \rangle) - \theta\|_{\mathbb{T}}^2 \leq \frac{C\|\varepsilon\|^2}{\tau^2 P^2},$$

where $P = \min_{(\lambda_1, \lambda_2) \in E''} |\langle x, \pi(\lambda_1)g \rangle \overline{\langle x, \pi(\lambda_2)g \rangle} + \varepsilon_{\lambda_1 \lambda_2}|$ and C is a universal constant.

ALGORITHM 4 (ANGULAR SYNCHRONIZATION)

Input: Graph $G'' = (\Lambda'', E'')$ with weighted edges.
Set A to be a weighted adjacency matrix defined in (9);
set $D = \text{diag}(d_\lambda)_{\lambda \in \Lambda''}$, where d_λ is the degree of the vertex λ ;
compute $L_1 = I - D^{-1/2}\bar{A}D^{-1/2}$;
compute the eigenvector u corresponding to the smallest eigenvalue α_1 of L_1 ;
set $\tilde{v}_\lambda = \frac{u_\lambda}{|u_\lambda|}$, $\lambda \in \Lambda$.
Output: The vector \tilde{v} approximating the vector of real phases of frame coefficients.

As Theorem 4.5 shows, the accuracy of Algorithm 4 depends on the spectral gap of the graph G'' . To find a subgraph $G'' \subset G'$ with spectral gap bounded away from zero and with the number of vertices satisfying $|\Lambda''| \geq M$, we shall use the *spectral clustering* algorithm [28, 1].

ALGORITHM 3 (SPECTRAL CLUSTERING)

Input: Graph $G' = (V', E')$, parameter τ .
While $\text{spg}(G) < \tau$ Set $D = \text{diag}(d_1, \dots, d_n)$, where d_i is the degree of the i th vertex;
set A to be the adjacency matrix of G' ;
compute the Laplacian $L = I - D^{-1/2}AD^{-1/2}$;
compute the eigenvector u of L with the second smallest eigenvalue α_2 ;
For $i = 1$ **to** $|V|/2$ let S_i be a set of vertices corresponding to i smallest entries of $D^{-1/2}u$;
set $h_i = \frac{|E(S_i, S_i^C)|}{\min\{\sum_{v \in S_i} \deg v, \sum_{v \in S_i^C} \deg v\}}$;
end for;
set $G' = G' \setminus S$, where $S = S_i$ with h_i minimal;
Output: Graph G with spectral gap at least τ .

To show that $|\Lambda''| \geq M$ in the graph G'' obtained after applying Algorithm 3, we use the following result. Its proof is based on the Cheeger inequality for the graph connection Laplacian [7] and can be found in [1].

Theorem 4.6. *Take $p \geq q \geq \frac{2}{3}$. Consider a regular graph $G = (V, E)$ with spectral gap $\lambda_2 > g(p, q) = 1 - 2(q(1 - q) - (1 - p))$ and set $\tau = \frac{1}{8}(\lambda_2 - g(p, q))^2$. Then, after Algorithm 2 removes at most $(1 - p)|V|$ vertices from G , Algorithm 3 outputs a subgraph with at least $q|V|$ vertices and spectral gap at least τ .*

We summarize the above discussion in the following reconstruction algorithm.

ALGORITHM 5 (RECONSTRUCTION IN THE NOISY CASE)

Input: Phaseless noisy measurements b w.r.t. $\Phi_\Lambda \cup \Phi_E$, given by (8); $F < \mathbb{Z}_M$, set $C \subset \mathbb{Z}_M$ and window g ; parameters α, β, τ .

Reconstruction: Construct graph $G = (\Lambda, E)$ with $\Lambda = F \times \mathbb{Z}_M$ and E as in (5);
assign to each $\lambda \in \Lambda$ weight b_λ ;
assign to each edge $(\lambda_1, \lambda_2) \in E$ weight $A_{\lambda_1 \lambda_2}$, given in (9);
run Algorithm 2 with parameters α, β to obtain $G' = (\Lambda', E')$;
run Algorithm 3 with parameter τ to obtain $G'' = (\Lambda'', E'')$;
run Algorithm 4 to obtain approximate phases $\{u_\lambda\}_{\lambda \in \Lambda''}$;
set $c_\lambda = u_\lambda \sqrt{b_\lambda}$, $\lambda \in \Lambda''$;
reconstruct $\tilde{x} = (\Phi_{\Lambda''} \Phi_{\Lambda''}^*)^{-1} \Phi_{\Lambda''} c$ from $c = \{c_\lambda\}_{\lambda \in \Lambda''}$.

Output: Approximation \tilde{x} of the initial signal x (up to a global phase).

Now we are ready to prove our main result.

Theorem 4.7. *Fix $x \in \mathbb{C}^M$ and consider the measurement procedure (8) described above, with $|F|$ and d sufficiently large. If the noise vector satisfies $\frac{\|\nu\|_2}{\|x\|_2} \leq \frac{C_1}{M}$ for some C_1 small enough, then there exists a constant C'' such that the estimate \tilde{x} produced by Algorithm 5 from the noisy measurements $\{b_j\}_{j=1}^N$ satisfies with overwhelming probability*

$$\min_{\theta \in [0, 2\pi)} \|\tilde{x} - e^{i\theta} x\|_2^2 \leq \frac{C'' \sqrt{M} \|\nu\|_2}{\Delta},$$

where $\Delta = \min_{\Lambda'' \subset \Lambda, |\Lambda''| \geq 2/3|\Lambda|} \sigma_{\min}^2(\Phi_{\Lambda''}^*)$.

Proof. Without loss of generality we can assume that $\|x\|_2 = 1$. As follows from Lemma 3.4, with overwhelming probability $\text{spg}(G) \geq 1 - \frac{6}{\sqrt{b}}$. Let us fix parameters $\tau_0 > 0$, $\alpha, \beta \in (0, 1)$ and apply Theorem 4.6 with $g(p, q) = 1 - 2(q(1 - q) - (1 - p)) = 1 - \frac{6}{\sqrt{b}} - \tau_0 < \text{spg}(G)$. Then, after Algorithm 2 deletes $(1 - p)|\Lambda| = (1 - \alpha - \beta)|\Lambda|$ vertices with the smallest and the largest corresponding measurements, we apply Algorithm 3 with parameter $\tau = \frac{1}{8}(\text{spg}(G) - g(p, q))^2 \geq \frac{\tau_0^2}{8}$. We obtain a graph $G'' = (V'', E'')$ with $|V''| \geq q|\Lambda|$ and $\text{spg}(G'') \geq \frac{\tau_0^2}{8}$.

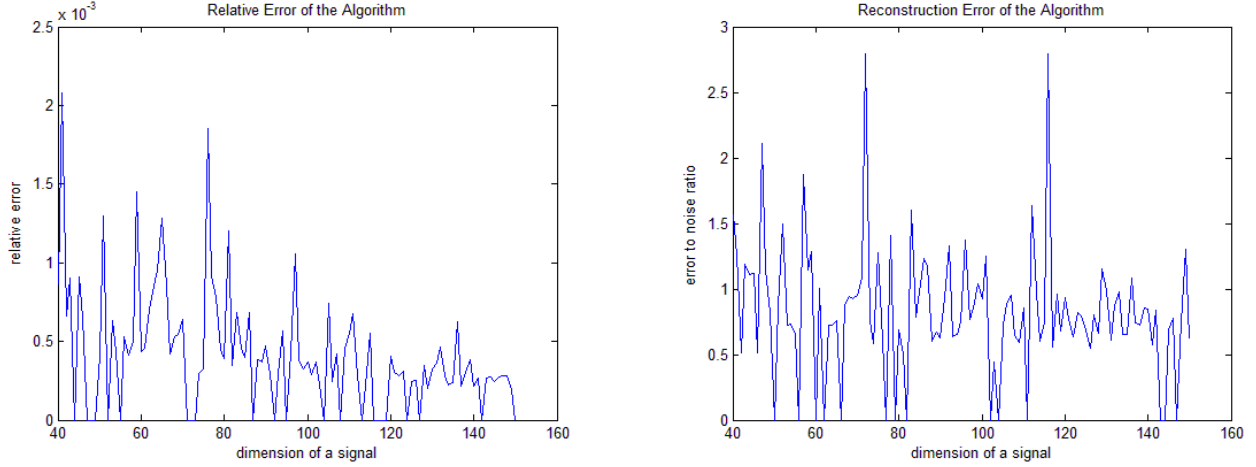


Figure 4: Dependence of the reconstruction error of the Algorithm 5 (left) and the error to noise ratio (right) on the signal dimension M . Here noise vector had independent normally distributed entries with variance $\sigma = 10^{-3}$.

Let us specify q now. Since we set $1 - 2(q(1 - q) - (1 - p)) = 1 - \frac{6}{\sqrt{b}} - \tau_0$, it follows $q(1 - q) = \frac{3}{\sqrt{b}} + (1 - \alpha - \beta) + \frac{\tau_0}{2}$. If τ_0 , α , β , and b are chosen appropriately, we can make $\frac{3}{\sqrt{b}} + (1 - \alpha - \beta) + \frac{\tau_0}{2} = A \leq \frac{2}{9}$. Then $1 \geq q = \frac{1 + \sqrt{1 - 4A}}{2} \geq \frac{1 + \sqrt{1/9}}{2} = \frac{2}{3}$. This ensures that $q \in (\frac{2}{3}, 1)$.

Now, after applying Algorithm 4 and using Theorem 4.5, we obtain a universal constant $C > 0$ and phase $\theta \in \mathbb{R}/2\pi\mathbb{Z}$, such that

$$\sum_{\lambda \in V''} \|\arg(u_\lambda) - \arg(\langle x, \pi(\lambda)g \rangle) - \theta\|_{\mathbb{T}}^2 \leq \frac{C\|\varepsilon\|_2^2}{\tau^2 P^2},$$

where $\varepsilon_{\lambda_1 \lambda_2} = \frac{1}{3} \sum_{t=0}^2 \omega^t \nu_{\lambda_1 \lambda_2 t}$. Since $\|\nu\|_2 \leq \frac{C_1}{M}$, we also have $|\nu_{\lambda_1 \lambda_2 t}| \leq \frac{C_1}{M}$, for all $\lambda_1, \lambda_2 \in \Lambda''$ and $t \in \{0, 1, 2\}$, and thus $|\varepsilon_{\lambda_1 \lambda_2}| \leq \frac{C_1}{M}$. By the Cauchy-Schwarz inequality we have

$$\|\varepsilon\|_2^2 = \sum_{(\lambda_1, \lambda_2) \in E''} |\varepsilon_{\lambda_1 \lambda_2}|^2 \leq \frac{1}{9} \sum_{(\lambda_1, \lambda_2) \in E''} \left(\sum_{i=0}^2 |\omega^i|^2 \right) \left(\sum_{i=0}^2 |\nu_{\lambda_1 \lambda_2 i}|^2 \right) \leq \frac{1}{3} \|\nu\|_2^2.$$

Theorem 4.1 implies that, for any $k, c > 0$ and $\epsilon = c^2$,

$$\left| \left\{ \lambda \in \Lambda, \text{ s.t. } |\langle x, \pi(\lambda)g \rangle| < \frac{c}{\sqrt{M}} \right\} \right| < |\Lambda|(\epsilon + k\sqrt{2\epsilon})$$

with probability at least $1 - 1/k^2$. Then, since $|\nu_\lambda| \leq \frac{C_1}{M}$, on this event we have

$$\left\{ \lambda \in \Lambda, \text{ s.t. } |\langle x, \pi(\lambda)g \rangle|^2 + \nu_\lambda < \frac{c^2}{M} - \frac{C_1}{M} \right\} \subset \left\{ \lambda \in \Lambda, \text{ s.t. } |\langle x, \pi(\lambda)g \rangle|^2 < \frac{c^2}{M} \right\},$$

that is,

$$\left| \left\{ \lambda \in \Lambda, \text{ s.t. } |\langle x, \pi(\lambda)g \rangle|^2 + \nu_\lambda < \frac{c^2 - C_1}{M} \right\} \right| < |\Lambda|(\epsilon + k\sqrt{2\epsilon}).$$

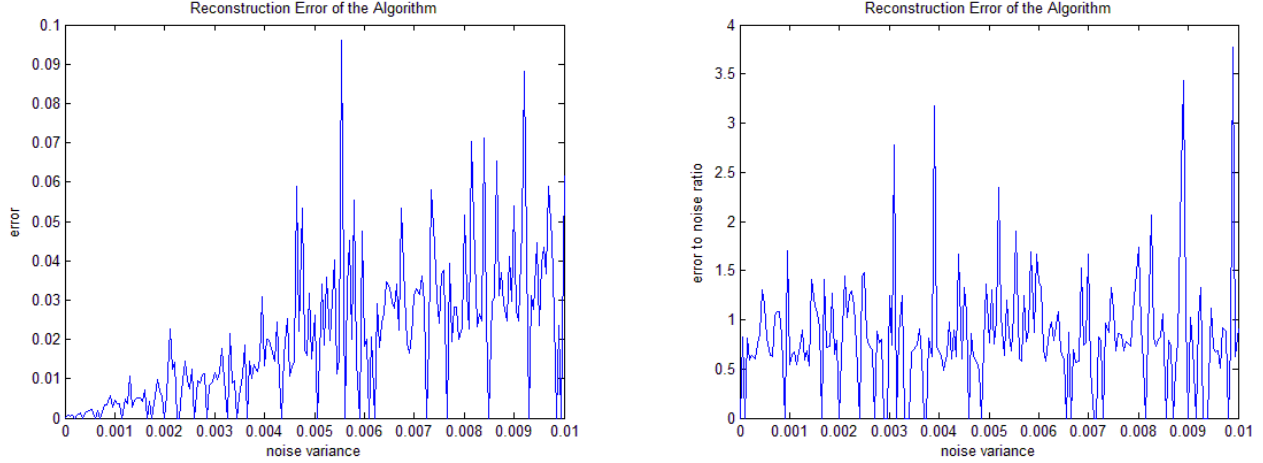


Figure 5: Dependence of the reconstruction error of Algorithm 5 (left) and the error to noise ratio (right) on the variance σ of the entries of the noise vector which are selected independently from $\mathcal{CN}(0, \sigma)$, for fixed signal dimension $M = 100$.

We set $\alpha = 1 - (\epsilon + k\sqrt{2\epsilon})$ and delete $|\Lambda|(\epsilon + k\sqrt{2\epsilon})$ vertices with the smallest corresponding measurements. For the remaining coefficients we have $|\langle x, \pi(\lambda)g \rangle| \geq \frac{\tilde{c}}{\sqrt{M}}$ for some constant $\tilde{c} > 0$, provided c^2 is sufficiently larger than C_1 . Similarly, setting $\beta = 1 - (\eta + k\sqrt{2\eta})$, we delete $|\Lambda|(\eta + k\sqrt{2\eta})$ vertices with the largest corresponding measurements, and for the remaining vertices, with probability at least $1 - \frac{1}{k^2}$, we have $|\langle x, \pi(\lambda)g \rangle| \leq \frac{\tilde{K}}{\sqrt{M}}$, for some constant $\tilde{K} > 0$.

Now, with probability at least $1 - \frac{2}{k^2}$, after applying Algorithms 1 and 2, we have $\frac{\tilde{c}}{\sqrt{M}} \leq |\langle x, \pi(\lambda)g \rangle| \leq \frac{\tilde{K}}{\sqrt{M}}$ for all $\lambda \in V''$. Thus

$$P = \min_{(\lambda_1, \lambda_2) \in E''} |\overline{\langle x, \pi(\lambda_1)g \rangle} \langle x, \pi(\lambda_2)g \rangle + \varepsilon_{\lambda_1 \lambda_2}| \geq \left| \frac{\tilde{c}^2}{M} - \max_{(\lambda_1, \lambda_2) \in E''} |\varepsilon_{\lambda_1 \lambda_2}| \right| \geq \left| \frac{\tilde{c}^2}{M} - \frac{C_1}{M} \right| \geq \frac{\tilde{C}}{M}.$$

Summing up, we obtain

$$\sum_{\lambda \in V''} \|\arg(u_\lambda) - \arg(\langle x, \pi(\lambda)g \rangle) - \theta\|_{\mathbb{T}}^2 \leq \frac{64C\|\nu\|_2^2 M^2}{3\tau_0^4 \tilde{C}^2}.$$

For every $\lambda \in V''$, we denote the obtained estimate of the corresponding frame coefficient by $c_\lambda = u_\lambda \sqrt{b_\lambda} = u_\lambda \sqrt{|\langle x, \pi(\lambda)g \rangle|^2 + \nu_\lambda}$. Also, set $\delta_\lambda = c_\lambda - e^{i\phi} \langle x, \pi(\lambda)g \rangle$ and $\zeta_\lambda = \sqrt{b_\lambda} - |\langle x, \pi(\lambda)g \rangle|$. Then

$$\begin{aligned} |\delta_\lambda| &= \left| \sqrt{b_\lambda} e^{i \arg(u_\lambda)} - \sqrt{b_\lambda} e^{i(\theta + \arg(\langle x, \pi(\lambda)g \rangle))} + \zeta_\lambda e^{i(\theta + \arg(\langle x, \pi(\lambda)g \rangle))} \right| \\ &\leq \sqrt{b_\lambda} |e^{i(\arg(u_\lambda) - \arg(\langle x, \pi(\lambda)g \rangle) - \theta)} - 1| + |\zeta_\lambda| \leq \sqrt{b_\lambda} \|\arg(u_\lambda) - \arg(\langle x, \pi(\lambda)g \rangle) - \theta\|_{\mathbb{T}} + |\zeta_\lambda|. \end{aligned}$$

Further, since $|\delta_\lambda|^2 \leq 2b_\lambda \|\arg(u_\lambda) - \arg(\langle x, \pi(\lambda)g \rangle) - \theta\|_{\mathbb{T}}^2 + 2|\zeta_\lambda|^2$, we obtain

$$\|\delta\|_2^2 \leq 2 \sum_{\lambda \in V''} b_\lambda \|\arg(u_\lambda) - \arg(\langle x, \pi(\lambda)g \rangle) - \theta\|_{\mathbb{T}}^2 + 2 \sum_{\lambda \in V''} |\zeta_\lambda|^2.$$

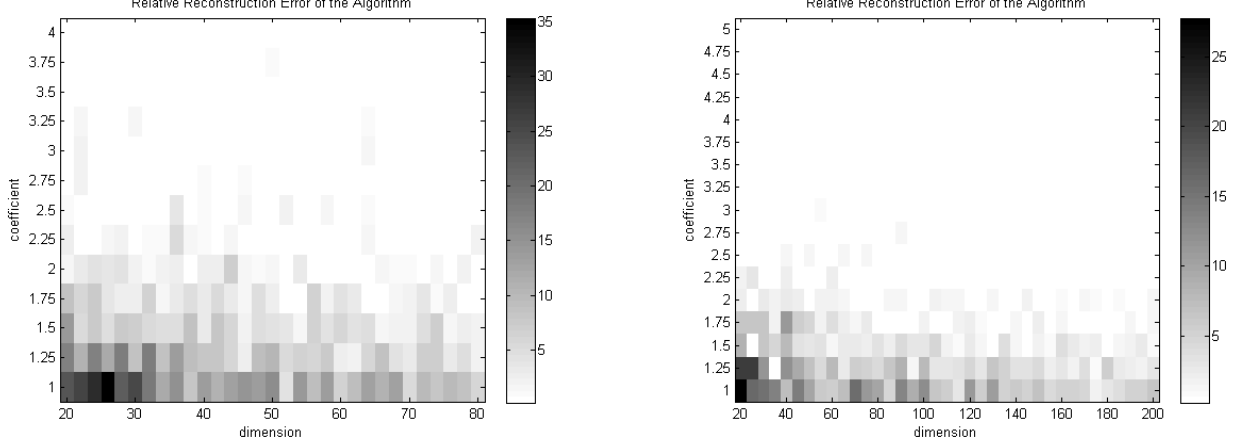


Figure 6: Dependence of the reconstruction error to noise ratio of Algorithm 5 on the parameter d for various dimensions. Parameter d controls the connectivity properties of the graph of measurements.

Since $(a - b)^2 \leq |a^2 - b^2|$ for all $a, b \in \mathbb{R}$, we obtain $\zeta_\lambda^2 = (\sqrt{b_\lambda} - |\langle x, \pi(\lambda)g \rangle|)^2 \leq |\nu_\lambda|$, and, since $b_\lambda \leq \frac{\tilde{K}^2}{M}$ and $\|\cdot\|_1 \leq \sqrt{|\Lambda''|} \|\cdot\|_2$,

$$\|\delta\|_2^2 \leq \frac{2\tilde{K}^2}{M} \sum_{\lambda \in V''} \|\arg(u_\lambda) - \arg(\langle x, \pi(\lambda)g \rangle) - \theta\|_{\mathbb{T}}^2 + 2\|\nu_V\|_1 \leq \frac{128C\tilde{K}^2\|\nu\|_2^2 M}{3\tau_0^4 \tilde{C}^2} + 2\sqrt{|F|M}\|\nu\|_2.$$

Since $\|\nu\|_2 \leq \frac{C_1}{M}$, we obtain

$$\|\delta\|_2^2 \leq C''\sqrt{M}\|\nu\|_2.$$

For the estimate \tilde{x} of x , constructed by Algorithm 5, we have

$$\tilde{x} = (\Phi_{\Lambda''}\Phi_{\Lambda''}^*)^{-1}\Phi_{\Lambda''}c = e^{i\theta}x + (\Phi_{\Lambda''}\Phi_{\Lambda''}^*)^{-1}\Phi_{\Lambda''}\delta.$$

As such, we obtain the desired bound on the reconstruction error

$$\|\tilde{x} - e^{i\theta}x\|_2^2 = \|(\Phi_{\Lambda''}\Phi_{\Lambda''}^*)^{-1}\Phi_{\Lambda''}\delta\|_2^2 \leq \|(\Phi_{\Lambda''}\Phi_{\Lambda''}^*)^{-1}\Phi_{\Lambda''}\|_2^2 \|\delta\|_2^2 = \frac{\|\delta\|_2^2}{\sigma_{\min}^2(\Phi_{\Lambda''}^*)} \leq \frac{C''\sqrt{M}\|\nu\|_2}{\sigma_{\min}^2(\Phi_{\Lambda''}^*)}.$$

□

5 Numerical results on the robustness of Algorithm 5

In the previous sections, we derived a measurement procedure and phase retrieval algorithms for noiseless and noisy time-frequency structured measurements. This section presents results from numerical simulations that illustrate how well our reconstruction algorithm performs in practice.

To support Theorem 4.7, we use two different sets of simulations. For the first kind, we let the dimension of the signal vary and explore the reconstruction error of the algorithm for a random normally distributed noise vector with independent entries and fixed variance. On

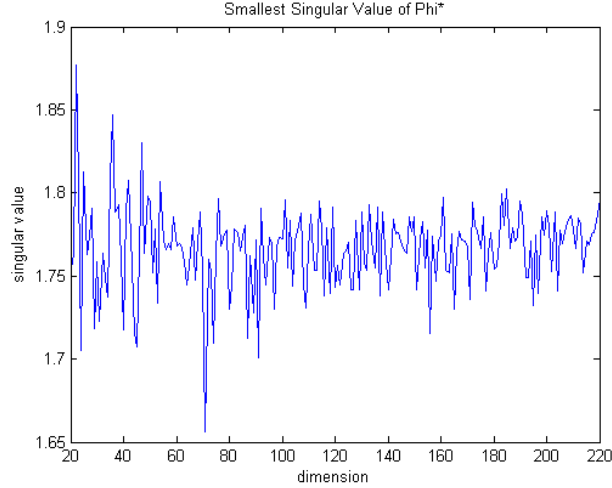


Figure 7: Dependence of the parameter $\Delta = \min_{\Lambda'' \subset \Lambda, |\Lambda''| \geq 2/3|\Lambda|} \sigma_{\min}^2(\Phi_{\Lambda''}^*)$ on the ambient dimension.

Figure 4 we show respective results. These suggest that the error to noise ratio does not depend on the signal dimension, unlike the bound obtained in Theorem 4.7, and is bounded above by a numerical constant close to 3. This suggests that the theoretical bound of ratio between reconstruction error of the algorithm and norm of the noise vector can be improved.

For the second simulation, we explore the dependence of reconstruction error and error to noise ratio on the noise variance for a fixed dimension and random, normally distributed noise vector with independent entries. Figure 5 shows the obtained results. These plots suggest that the reconstruction error grows linearly with the magnitude of noise. One of the main goals for future research is to understand the gap between theoretically predicted robustness guarantees and results obtained numerically (Figures 4 and 5) and improve robustness guarantees for the reconstruction algorithm, so that they better reflect the actual behavior.

The bound of the reconstruction error of Algorithm 5 obtained in Theorem 4.7 depends on several parameters, including the spectral gap τ of the graph G and the parameter

$$\Delta = \min_{\Lambda'' \subset \Lambda, |\Lambda''| \geq 2/3|\Lambda|} \sigma_{\min}^2(\Phi_{\Lambda''}^*).$$

We now investigate these dependencies as well as the behavior of the above mentioned parameters numerically.

We start with the spectral gap τ of the graph $G = (\Lambda, E)$. As follows from Lemma 3.4, τ depends on the parameter d in the construction of the set C in (5). Numerical results presented on Figure 6 show the dependence of the reconstruction error to noise ratio on the constant d (vertical axes) for various dimensions (horizontal axis). One can see that starting approximately at $d = 3$, that is, much earlier than a theoretically predicted value $d = 144$, this ratio does not exceed 4.

We also investigate the parameter Δ , which is the lower bound on the smallest singular values for all submatrices $\Phi_{\Lambda''}^*$ of Φ_{Λ}^* , with $|\Lambda''| \geq 2/3|\Lambda|$. To the best of our knowledge, there are no results available on the smallest singular value of the analysis map Φ^* of a Gabor frame with a random window and general Λ with cardinality $|\Lambda| = O(M)$. We tested the value Δ numerically, the obtained results are presented on Figure 7. This numerical results show that Δ

is bounded away from zero by a numerical constant not depending on the dimension M . Proving this conjectured bound is another important task for our future research.

Acknowledgments

P. Salanevich thanks Prof. Dr. Felix Krahmer for insightful discussions during the “Mathematics of Signal Processing” trimester program (Hausdorff Research Institute for Mathematics, Bonn, Germany), and Prof. Dr. Terence Tao for his valuable suggestions during her visit at UCLA in Spring 2015. P. Salanevich also thanks Prof. Dr. Holger Boche for helpful discussions during her visit to Technical University of Munich in December 2014. P. Salanevich is supported by Research Grant for Doctoral Candidates and Young Academics and Scientists of German Academic Exchange Service (DAAD), the funding is greatly appreciated.

Appendix: proof of Theorem 4.1

Here we prove Theorem 4.1 which is formulated as follows.

Theorem. *Fix $x \in \mathbb{S}^{M-1} \subset \mathbb{C}^M$ and consider a Gabor frame (g, Λ) with $\Lambda \subset \mathbb{Z}_M \times \mathbb{Z}_M$ and a random window g uniformly distributed on the unit sphere \mathbb{S}^{M-1} . Then the following holds.*

(a) *For any $c > 0$ and $k > 0$, with probability at least $1 - \frac{1}{k^2}$, we have*

$$\left| \left\{ \lambda \in \Lambda, \text{ s.t. } |\langle x, \pi(\lambda)g \rangle| < \frac{c}{\sqrt{M}} \right\} \right| < |\Lambda|(c^2 + kc).$$

(b) *For any $K > 0$ and $k > 0$, with probability at least $1 - \frac{1}{k^2}$, we have*

$$\left| \left\{ \lambda \in \Lambda, \text{ s.t. } |\langle x, \pi(\lambda)g \rangle| > \frac{K}{\sqrt{M}} \right\} \right| < |\Lambda| \left(\frac{8}{\pi} e^{-K^2} + k \frac{2\sqrt{2}}{\sqrt{\pi}} e^{-\frac{K^2}{2}} \right).$$

Proof. (a) For $x \in \mathbb{S}^{M-1}$ fixed, we set $G_\delta(x) = \{\phi \in \mathbb{S}^{M-1} \subset \mathbb{C}^M, \text{ s.t. } |\langle x, \phi \rangle| < \delta\}$. We are interested in the distribution of the random variable

$$Z_x = |G_\delta(x) \cap (g, \Lambda)| = \sum_{\lambda \in \Lambda} \mathbf{1}_{\{\pi(\lambda)g \in G_\delta(x)\}},$$

where $\mathbf{1}_{\{\phi \in G_\delta(x)\}}$ is the characteristic function of the event $\{\phi \in G_\delta(x)\}$. In words, Z_x is the number of small measurements of the fixed signal x with respect to the Gabor frame (g, Λ) with a random window g .

First, note that for each $\lambda \in \Lambda$, $\pi(\lambda)g$ is also uniformly distributed on \mathbb{S}^{M-1} , that is, the random vectors $\pi(\lambda)g$, $\lambda \in \Lambda$, are equally distributed. Indeed, consider a random vector h , such that $h(m) \sim i.i.d. \mathcal{CN}(0, \frac{1}{M})$. Then it is well known that $h/\|h\|_2$ is uniformly distributed on \mathbb{S}^{M-1} , since random vector $h/\|h\|_2$ almost surely has unit norm and its distribution is rotation invariant [25]. In other words, we can write $g = h/\|h\|_2$. As such, we have $\pi(\lambda)g = \pi(\lambda)h/\|h\|_2 = \pi(\lambda)h/\|\pi(\lambda)h\|_2$, since both modulation and time shift are unitary operators. The coordinates

of h are independent identically distributed random variables, thus translation T_k , which is just a permutation of the vector coordinates, preserves the distribution of h . For the modulation we have

$M_\ell h(m) = e^{2\pi i m \ell / M} h(m)$ is also normally distributed, with

$$\mathbb{E}(e^{2\pi i m \ell / M} h(m)) = e^{2\pi i m \ell / M} \mathbb{E}(h(m)) = 0;$$

$$\text{Var}(e^{2\pi i m \ell / M} h(m)) = \mathbb{E}(e^{2\pi i m \ell / M} h(m) e^{-2\pi i m \ell / M} \bar{h}(m)) = \mathbb{E}(h(m) \bar{h}(m)) = \text{Var}(h(m)) = \frac{1}{M}.$$

Thus the distribution of h is preserved by both modulation and time shift, and $\pi(\lambda)g = \pi(\lambda)h/||\pi(\lambda)h||$ has the same distribution as $g = h/||h||_2$.

Since $\pi(\lambda)g$ has the same distribution as g , we have $\mathbb{P}\{|\langle x, \pi(\lambda)g \rangle| < \delta\} = \mathbb{P}\{|\langle x, g \rangle| < \delta\}$, for all $\lambda \in \Lambda$. Thus

$$\mathbb{E}(Z_x) = \mathbb{E}\left(\sum_{\lambda \in \Lambda} \mathbf{1}_{\{\pi(\lambda)g \in G_\delta(x)\}}\right) = \sum_{\lambda \in \Lambda} \mathbb{P}\{\pi(\lambda)g \in G_\delta(x)\} = |\Lambda| \mathbb{P}\{|\langle x, g \rangle| < \delta\}. \quad (11)$$

Let R be an orthogonal matrix, such that $x = R e_1$, where $e_1 = (1, 0, 0, \dots, 0)^T$ is the first vector of the standard basis. Then $|\langle x, g \rangle| = |\langle R e_1, g \rangle| = |\langle e_1, R^* g \rangle|$. By the rotational symmetry of the distribution of g , we obtain

$$\mathbb{P}\{|\langle x, g \rangle| < \delta\} = \mathbb{P}\{|\langle e_1, g \rangle| < \delta\} = \mathbb{P}\{|g(0)| < \delta\}.$$

Let us identify the complex unit sphere $\mathbb{S}^{M-1} \subset \mathbb{C}^M$ with the real unit sphere $\mathbb{S}_{\mathbb{R}}^{2M-1} \subset \mathbb{R}^{2M}$, using the map $\mathcal{I} : \mathbb{S}^{M-1} \rightarrow \mathbb{S}_{\mathbb{R}}^{2M-1}$ given by

$$\mathcal{I}(z_0, z_1, \dots, z_{M-1}) = (\Re(z_0), \Im(z_0), \Re(z_1), \Im(z_1), \dots, \Re(z_{M-1}), \Im(z_{M-1})). \quad (12)$$

Since g is uniformly distributed on \mathbb{S}^{M-1} , $\tilde{g} = \mathcal{I}(g)$ is uniformly distributed on $\mathbb{S}_{\mathbb{R}}^{2M-1}$. Thus

$$\mathbb{P}\{|g(0)| < \delta\} = \mathbb{P}\{\tilde{g}(0)^2 + \tilde{g}(1)^2 < \delta^2\} = \frac{S_{<\delta}}{S_1},$$

where $S_1 = \frac{2\pi^M}{(M-1)!}$ is the surface area of $\mathbb{S}_{\mathbb{R}}^{2M-1}$, and $S_{<\delta}$ is the surface area of the set $\{z \in \mathbb{S}_{\mathbb{R}}^{2M-1}, \text{ s.t. } z_0^2 + z_1^2 < \delta^2\}$.

$$S_{<\delta} = \int_{-\delta}^{\delta} \int_{-\sqrt{\delta^2 - z_0^2}}^{\sqrt{\delta^2 - z_0^2}} \frac{2\pi^{M-1}}{(M-2)!} (1 - z_0^2 - z_1^2)^{\frac{2M-1}{2}} dz_1 dz_0 < \frac{2\pi^M \delta^2}{(M-2)!},$$

that is, $\mathbb{P}\{|g(0)| < \delta\} \leq \delta^2(M-1)$. Now, setting $\delta = \frac{c}{\sqrt{M}}$ for c sufficiently small, we obtain

$$\mathbb{P}\left\{|\langle x, g \rangle| < \frac{c}{\sqrt{M}}\right\} \leq c^2. \quad (13)$$

Using equations (11) and (13), we obtain

$$\mu = \mathbb{E}(Z_x) = |\Lambda| \mathbb{P}\left\{|\langle x, g \rangle| < \frac{c}{\sqrt{M}}\right\} \leq |\Lambda| c^2. \quad (14)$$

Similarly, using (13), for the variance of Z_x we obtain

$$\begin{aligned}
\sigma^2 &= \text{Var}(Z_x) = \mathbb{E}(Z_x^2) - (\mathbb{E}(Z_x))^2 \leq \mathbb{E}(Z_x^2) = \mathbb{E} \left(\left(\sum_{\lambda \in \Lambda} \mathbf{1}_{\{\pi(\lambda)g \in G_{\frac{c}{\sqrt{M}}}(x)\}} \right)^2 \right) = \\
&= \mathbb{E} \left(\sum_{\lambda \in \Lambda} \mathbf{1}_{\{\pi(\lambda)g \in G_{\frac{c}{\sqrt{M}}}(x)\}}^2 + \sum_{\substack{(\lambda_1, \lambda_2) \in \Lambda \times \Lambda, \\ \lambda_1 \neq \lambda_2}} \mathbf{1}_{\{\pi(\lambda_1)g \in G_{\frac{c}{\sqrt{M}}}(x)\}} \mathbf{1}_{\{\pi(\lambda_2)g \in G_{\frac{c}{\sqrt{M}}}(x)\}} \right) = \\
&= \sum_{\lambda \in \Lambda} \mathbb{P} \left\{ \pi(\lambda)g \in G_{\frac{c}{\sqrt{M}}}(x) \right\} + \sum_{\substack{(\lambda_1, \lambda_2) \in \Lambda \times \Lambda, \\ \lambda_1 \neq \lambda_2}} \mathbb{P} \left\{ \pi(\lambda_1)g \in G_{\frac{c}{\sqrt{M}}}(x) \text{ and } \pi(\lambda_2)g \in G_{\frac{c}{\sqrt{M}}}(x) \right\} \leq \\
&\leq |\Lambda| \mathbb{P} \left\{ |\langle x, g \rangle| < \frac{c}{\sqrt{M}} \right\} + (|\Lambda|^2 - |\Lambda|) \mathbb{P} \left\{ |\langle x, g \rangle| < \frac{c}{\sqrt{M}} \right\} \leq c^2 |\Lambda|^2, \tag{15}
\end{aligned}$$

that is, $\sigma \leq c|\Lambda|$. Then, using Chebychev inequality and bounds (14) and (15), we have

$$\mathbb{P}\{Z_x \geq |\Lambda|(c^2 + kc)\} \leq \mathbb{P}\{Z_x \geq \mu + k\sigma\} \leq \mathbb{P}\{|Z_x - \mu| \geq k\sigma\} \leq \frac{1}{k^2},$$

for any $k > 0$. In other words, if we delete $|\Lambda|(c^2 + kc)$ smallest phaseless measurements, for the remaining measurements we would have $|\langle x, \pi(\lambda)g \rangle| \geq \frac{c}{\sqrt{M}}$ with probability at least $1 - \frac{1}{k^2}$. This concludes the proof of part (a).

The proof of (b) follows the same steps. Let K be a constant and consider the following random variable:

$$U_x = \sum_{\lambda \in \Lambda} \mathbf{1}_{\{|\langle x, \pi(\lambda)g \rangle| > K/\sqrt{M}\}}.$$

Since, for each $\lambda \in \Lambda$, $\pi(\lambda)g$ has the same (uniform on \mathbb{S}^{M-1}) distribution as g , we have $\mathbb{P}\{|\langle x, \pi(\lambda)g \rangle| > K/\sqrt{M}\} = \mathbb{P}\{|\langle x, g \rangle| > K/\sqrt{M}\}$, for all $\lambda \in \Lambda$. As above, we have

$$\mathbb{P}\{|\langle x, g \rangle| > K/\sqrt{M}\} = \mathbb{P}\{|\langle e_1, g \rangle| > K/\sqrt{M}\} = \mathbb{P}\{|g(0)| > K/\sqrt{M}\}.$$

Using the map $\mathcal{I} : \mathbb{S}^{M-1} \rightarrow \mathbb{S}_{\mathbb{R}}^{2M-1}$ defined in (12), for $\tilde{g} = \mathcal{I}(g)$ we obtain

$$\mathbb{P} \left\{ |g(0)| > \frac{K}{\sqrt{M}} \right\} = \mathbb{P} \left\{ \tilde{g}(0)^2 + \tilde{g}(1)^2 > \frac{K^2}{M} \right\} = \frac{S_{>K/\sqrt{M}}}{S_1},$$

where $S_1 = \frac{2\pi^M}{(M-1)!}$ is the surface area of $\mathbb{S}_{\mathbb{R}}^{2M-1}$, and $S_{>K/\sqrt{M}}$ is the surface area of the set

$$\left\{ z \in \mathbb{S}_{\mathbb{R}}^{2M-1}, \text{ s.t. } z_0^2 + z_1^2 > \frac{K^2}{M} \right\}.$$

$$\begin{aligned}
S_{>K/\sqrt{M}} &= \int_{|z_0| \leq 1} \int_{\sqrt{\frac{K^2}{M} - z_0^2} < |z_1| < \sqrt{1-z_0^2}} \frac{2\pi^{M-1}}{(M-2)!} (1 - z_0^2 - z_1^2)^{\frac{2M-1}{2}} dz_1 dz_0 \\
&= 8 \frac{2\pi^{M-1}}{(M-2)!} \left(\int_0^{\frac{K}{\sqrt{2M}}} \int_{\sqrt{\frac{K^2}{M} - z_0^2}}^{\sqrt{1-z_0^2}} (1 - z_0^2 - z_1^2)^{\frac{2M-1}{2}} dz_1 dz_0 + \int_{\frac{K}{\sqrt{2M}}}^{\frac{1}{\sqrt{2}}} \int_{z_0}^{\sqrt{1-z_0^2}} (1 - z_0^2 - z_1^2)^{\frac{2M-1}{2}} dz_1 dz_0 \right) \\
&\leq \frac{16\pi^{M-1}}{(M-2)!} \frac{\sqrt{2M}}{K} \left(\int_0^{\frac{K}{\sqrt{2M}}} \int_{\sqrt{\frac{K^2}{M} - z_0^2}}^{\sqrt{1-z_0^2}} z_1 (1 - z_0^2 - z_1^2)^{\frac{2M-1}{2}} dz_1 dz_0 + \int_{\frac{K}{\sqrt{2M}}}^{\frac{1}{\sqrt{2}}} \int_{z_0}^{\sqrt{1-z_0^2}} z_1 (1 - z_0^2 - z_1^2)^{\frac{2M-1}{2}} dz_1 dz_0 \right) \\
&= \frac{8\pi^{M-1}}{(M-2)!} \frac{\sqrt{2M}}{K} \frac{2}{2M+1} \left(\int_0^{\frac{K}{\sqrt{2M}}} \left((1 - z_0^2 - z_1^2)^{\frac{2M+1}{2}} \Big|_{\sqrt{1-z_0^2}}^{\sqrt{\frac{K^2}{M} - z_0^2}} \right) dz_0 + \int_{\frac{K}{\sqrt{2M}}}^{\frac{1}{\sqrt{2}}} \left((1 - z_0^2 - z_1^2)^{\frac{2M+1}{2}} \Big|_{\sqrt{1-z_0^2}}^{z_0} \right) dz_0 \right) \\
&= \frac{8\pi^{M-1}}{(M-2)!} \frac{\sqrt{2M}}{K} \frac{2}{2M+1} \left(\frac{K}{\sqrt{2M}} \left(1 - \frac{K^2}{M} \right)^{\frac{2M+1}{2}} + \int_{\frac{K}{\sqrt{2M}}}^{\frac{1}{\sqrt{2}}} (1 - 2z_0^2)^{\frac{2M+1}{2}} dz_0 \right) \\
&\leq \frac{8\pi^{M-1}}{(M-1)!} \left(\left(1 - \frac{K^2}{M} \right)^{\frac{2M+1}{2}} + \frac{M}{2K^2} \int_{\frac{K}{\sqrt{2M}}}^{\frac{1}{\sqrt{2}}} 4z_0 (1 - 2z_0^2)^{\frac{2M+1}{2}} dz_0 \right) \\
&\leq \frac{8\pi^{M-1}}{(M-1)!} \left(\left(1 - \frac{K^2}{M} \right)^{\frac{2M+1}{2}} + \frac{1}{2K^2} \left(1 - \frac{K^2}{M} \right)^{\frac{2M+3}{2}} \right) \leq \frac{8\pi^{M-1}}{(M-1)!} \left(e^{-\frac{K^2}{M} \frac{2M+1}{2}} + \frac{1}{2K^2} e^{-\frac{K^2}{M} \frac{2M+3}{2}} \right) \\
&\leq \frac{16\pi^{M-1}}{(M-1)!} e^{-K^2}.
\end{aligned}$$

Here, we used the symmetry of the domain of integration, the fact that $z_0^2 + z_1^2 < \frac{K^2}{M}$ implies $\max\{|z_0|, |z_1|\} > \frac{K}{\sqrt{2M}}$, and inequality $1 - x \leq e^{-x}$. Using the computed bound for $S_{>K/\sqrt{M}}$, we obtain $\mathbb{P}\{|g(0)| > \frac{K}{\sqrt{M}}\} \leq \frac{8}{\pi} e^{-K^2}$. Then

$$\mu = \mathbb{E}(U_x) = \sum_{\lambda \in \Lambda} \mathbb{P}\{|\langle x, \pi(\lambda)g \rangle| > K/\sqrt{M}\} = |\Lambda| \mathbb{P}\{|\langle x, g \rangle| > K/\sqrt{M}\} \leq |\Lambda| \frac{8}{\pi} e^{-K^2}. \quad (16)$$

Similarly, for the variance of U_x we obtain

$$\begin{aligned}
\sigma^2 = \text{Var}(U_x) &= \mathbb{E}(U_x^2) - (\mathbb{E}(U_x))^2 \leq \mathbb{E}(U_x^2) = \mathbb{E} \left(\left(\sum_{\lambda \in \Lambda} \mathbf{1}_{\{|\langle x, \pi(\lambda)g \rangle| > K/\sqrt{M}\}} \right)^2 \right) = \\
&= \sum_{\lambda \in \Lambda} \mathbb{P} \left\{ |\langle x, \pi(\lambda)g \rangle| > \frac{K}{\sqrt{M}} \right\} + \sum_{\substack{(\lambda_1, \lambda_2) \in \Lambda \times \Lambda, \\ \lambda_1 \neq \lambda_2}} \mathbb{P} \left\{ |\langle x, \pi(\lambda_1)g \rangle| > \frac{K}{\sqrt{M}}, |\langle x, \pi(\lambda_2)g \rangle| > \frac{K}{\sqrt{M}} \right\} \leq \\
&= |\Lambda| \mathbb{P}\{|\langle x, g \rangle| > K/\sqrt{M}\} + (|\Lambda|^2 - |\Lambda|) \mathbb{P}\{|\langle x, g \rangle| > K/\sqrt{M}\} \leq \frac{8}{\pi} e^{-K^2} |\Lambda|^2, \quad (17)
\end{aligned}$$

that is, $\sigma \leq \frac{2\sqrt{2}}{\sqrt{\pi}} e^{-\frac{K^2}{2}} |\Lambda|$. Then, using Chebychev inequality and bounds (16) and (17), we obtain

$$\mathbb{P} \left\{ U_x \geq |\Lambda| \left(\frac{8}{\pi} e^{-K^2} + k \frac{2\sqrt{2}}{\sqrt{\pi}} e^{-\frac{K^2}{2}} \right) \right\} \leq \mathbb{P}\{U_x \geq \mu + k\sigma\} \leq \mathbb{P}\{|U_x - \mu| \geq k\sigma\} \leq \frac{1}{k^2},$$

for any $k > 0$. In other words, if we delete $|\Lambda| \left(\frac{8}{\pi} e^{-K^2} + k \frac{2\sqrt{2}}{\sqrt{\pi}} e^{-\frac{K^2}{2}} \right)$ largest phaseless measurements, with probability at least $1 - \frac{1}{k^2}$ for the remaining measurements we would have $|\langle x, \pi(\lambda)g \rangle| \leq \frac{K}{\sqrt{M}}$. \square

6 References

- [1] Boris Alexeev, Afonso S. Bandeira, Matthew Fickus, and Dustin G. Mixon. Phase retrieval with polarization. *SIAM Journal on Imaging Sciences*, 7(1):35–66, 2014.
- [2] Boris Alexeev, Jameson Cahill, and Dustin G. Mixon. Full spark frames. *Journal of Fourier Analysis and Applications*, 18(6):1167–1194, 2012.
- [3] Noga Alon. Eigenvalues and expanders. *Combinatorica*, 6(2):83–96, 1986.
- [4] Radu Balan, Bernhard G. Bodmann, Peter G. Casazza, and Dan Edidin. Painless reconstruction from magnitudes of frame coefficients. *Journal of Fourier Analysis and Applications*, 15(4):pp 488–501, August 2009.
- [5] Radu Balana, Peter Casazza, and Dan Edidin. On signal reconstruction without phase. *Applied and Computational Harmonic Analysis*, 20(3):pp.345–356, May 2006.
- [6] Afonso S. Bandeira, Yutong Chen, and Dustin G. Mixon. Phase retrieval from power spectra of masked signals. *Information and Inference*, page iau002, 2014.
- [7] Afonso S. Bandeira, Amit Singer, and Daniel A. Spielman. A Cheeger inequality for the graph connection Laplacian. *SIAM Journal on Matrix Analysis and Applications*, 34(4):1611–1630, 2013.
- [8] Irena Bojarovska and Axel Flinth. Phase retrieval from gabor measurements. *Journal of Fourier Analysis and Applications*, 22(3):542–567, 2016.
- [9] Oliver Bunk, Ana Diaz, Franz Pfeiffer, Christian David, Bernd Schmitt, Dillip K. Satapathy, and J. Friso van der Veen. Diffractive imaging for periodic samples: retrieving one-dimensional concentration profiles across microfluidic channels. *Acta Crystallographica Section A: Foundations of Crystallography*, 63(4):306–314, 2007.
- [10] Emmanuel J. Candès, Yonina C. Eldar, Thomas Strohmer, and Vladislav Voroninski. Phase retrieval via matrix completion. *SIAM Journal on Imaging Sciences*, 6(1):pp 199–225, February 2013.
- [11] Emmanuel J. Candès and Xiaodong Li. Solving quadratic equations via PhaseLift when there are about as many equations as unknowns. *Foundations of Computational Mathematics*, pages pp1–10, 2012.
- [12] Emmanuel J. Candès, Xiaodong Li, and Mahdi Soltanolkotabi. Phase retrieval from coded diffraction patterns. *Applied and Computational Harmonic Analysis*, 39(2):277–299, 2015.

- [13] Emmanuel J. Candès, Xiaodong Li, and Mahdi Soltanolkotabi. Phase retrieval via Wirtinger flow: Theory and algorithms. *Information Theory, IEEE Transactions on*, 61(4):1985–2007, 2015.
- [14] Emmanuel J. Candès, Thomas Strohmer, and Vladislav Voroninski. PhaseLift: Exact and stable signal recovery from magnitude measurements via convex programming. *Communications on Pure and Applied Mathematics*, 66(8):pp 1241–1274, August 2013.
- [15] Yuxin Chen and Emmanuel J. Candès. Solving random quadratic systems of equations is nearly as easy as solving linear systems. In *Advances in Neural Information Processing Systems*, pages 739–747, 2015.
- [16] James R. Fienup and J. Christopher Dainty. Phase retrieval and image reconstruction for astronomy. pages 231–275. San Diego, CA, USA: Academic, 1987.
- [17] Ralph W Gerchberg. A practical algorithm for the determination of phase from image and diffraction plane pictures. *Optik*, 35:237, 1972.
- [18] Prahladh Harsha and Adam Barth. Lecture 5: Derandomization (Part II). Available at <http://www.tcs.tifr.res.in/~prahladh/teaching/05spring/lectures/lec5.pdf>.
- [19] Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43(4):pp.439–561, 2006.
- [20] Mark Iwen, Aditya Viswanathan, and Yang Wang. Fast phase retrieval from local correlation measurements. *arXiv preprint arXiv:1501.02377*, 2015.
- [21] Felix Krahmer, Shahar Mendelson, and Holger Rauhut. Suprema of chaos processes and the restricted isometry property. *Communications on Pure and Applied Mathematics*, 67(11):1877–1904, 2014.
- [22] James Lawrence, Götz E. Pfander, and David Walnut. Linear independence of Gabor systems in finite dimensional vector spaces. *Journal of Fourier Analysis and Applications*, 11(6):pp. 715–726, 2006.
- [23] Alexander Lubotzky, Ralph Phillips, and Peter Sarnak. Explicit expanders and the Ramanujan conjectures. In *Proceedings of the eighteenth annual ACM symposium on Theory of computing*, pages 240–246. ACM, 1986.
- [24] Romanos-Diogenes Malikiosis. A note on Gabor frames in finite dimensions. *Applied and Computational Harmonic Analysis*, 38(2):318–330, 2015.
- [25] George Marsaglia et al. Choosing a point from the surface of a sphere. *The Annals of Mathematical Statistics*, 43(2):645–646, 1972.
- [26] Jianwei Miao, Tetsuya Ishikawa, Qun Shen, and Thomas Earnest. Extending X-ray crystallography to allow the imaging of non-crystalline materials, cells and single protein complexes. *Annu. Rev. Phys. Chem.*, 59:pp.387–410, 2008.

- [27] Rick P. Millane. Phase retrieval in crystallography and optics. *JOSA A*, 7(3):394–411, 1990.
- [28] Andrew Y. Ng, Michael I. Jordan, Yair Weiss, et al. On spectral clustering: Analysis and an algorithm. *Advances in neural information processing systems*, 2:849–856, 2002.
- [29] Götz E. Pfander. Gabor frames in finite dimensions. In Peter G. Casazza and Gitta Kutyniok, editors, *Finite Frames: Theory and Applications*. Birkhäuser Boston, 2013.
- [30] Haralambos Sahinoglou and Sergio D. Cabrera. On phase retrieval of finite-length sequences using the initial time sample. *Circuits and Systems, IEEE Transactions on*, 38(8):954–958, 1991.
- [31] Palina Salanevich and Götz E. Pfander. Polarization based phase retrieval for time-frequency structured measurements. In *Sampling Theory and Applications (SampTA), 2015 International Conference on*, pages 187–191. IEEE, 2015.
- [32] Amit Singer. Angular synchronization by eigenvectors and semidefinite programming. *Applied and computational harmonic analysis*, 30(1):20–36, 2011.
- [33] Terence Tao and Van H. Vu. *Additive combinatorics*, volume 13. Cambridge University Press, 2006.
- [34] James D. Watson and Francis H. C. Crick. A structure for deoxyribose nucleic acid. *A century of Nature: twenty-one discoveries that changed science and the world*, page 82, 2003.
- [35] Avi Wigderson. Zigzag products, expander constructions, connections, and applications. *Lecture Notes in Computer Science*, pages 443–444, 2003.